# QTL ANALYSIS OF YIELD COMPONENTS IN COTTON

A Thesis

by

LORIN MICHAEL HARVEY

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

| | |
|---|---|
| Chair of Committee, | Steve Hague |
| Co-Chair of Committee, | Hongbin Zhang |
| Committee Members, | Wayne Smith |
| | Sheilah Oltmans-Deardorff |
| Head of Department, | David Baltensperger |

May 2016

Major Subject: Plant Breeding

## ABSTRACT

Cotton is the world's leading natural textile fiber and oilseed crop, with a global economic impact of over $600 billion. However, few QTLs have been identified and linked with yield traits. In this study, a large number of SNP markers were mapped across all 26 cotton chromosomes using an inter-specific RIL population. It is also one of the first genetic maps created using markers obtained by ddRADseq.

A total of 18 QTLs controlling lint percent and seed index were identified. The phenotypic effect of each of the QTLs ranges from 6.15 to 37.48%. However, only one of the QTLs was discovered in multiple environments. This suggests that the QTLs, and subsequently the genes controlling yield traits are highly affected by GxE interactions and gene-gene interactions. The goal of being able to apply MAS to wide spread geographical areas in an attempt to improve the efficiency of cotton breeding might not be practical. MAS may be better suited for small targeted areas, where environments are more similar.

## DEDICATION

This thesis is dedicated to my parents Roger and Carol Harvey whose help, love, and guidance throughout the years has helped me to achieve all of my goals. Without their support, I would not be who I am today.

I would also like to dedicate this thesis to my wife Rebekah Harvey whose undying love and support have helped me throughout the past year. Without you I would not be where I am today.

## ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Steve Hague, my committee co-chair Dr. Hongbin Zhang, and my committee members, Dr. C. Wayne Smith, and Dr. Sheilah Oltmans-Deardorff, for their guidance and support throughout the course of this research.

I would also like to take this time to thank Dawn Deno, fellow graduate students, and student workers at the Cotton Improvement Lab who have helped me to complete this project. Without their support and assistance, I would have never been able to complete this research.

Thanks to USDA/NIFA for financial support. Without their funding, this project would not have been possible.

Finally, thanks to my mother and father for their encouragement and to my wife for her patience, love, and support.

## NOMENCLATURE

| | |
|---|---|
| HVI | High Volume Instrumentation |
| MAS | Marker Assisted Selection |
| SNP | Single Nucleotide Polymorphism |
| QTL | Quantitative Trait Loci |
| SSR | Single Sequence Repeat |
| RAPD | Random Amplified Polymorphic DNA |
| AFLP | Amplified Fragment Length Polymorphism |
| RFLP | Restriction Fragment Length Polymorphism |
| PCR | Polymerase Chain Reaction |
| HVI | High Volume Instrument |
| GBS | Genotyping by Sequencing |

## TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

# CHAPTER I

## INTRODUCTION

Cotton (*Gossypium spp.*) is the world's leading natural textile fiber and oilseed crop. According to Wang K. et al., (2012), the global market value of textile was approximately $630.6 billion in 2011. According to the USDA (2016), global production of cotton seed was 43.1 million tons, with an estimated value of $8.53 billion in 2014. Upland cotton (*Gossypium hirsutum)* makes up 95% of the 33.1 million hectares planted to cotton world-wide (Johnson et al., 2014). However, the US share of the global cotton market has decreased over the last 10 years, and continued cotton research could help to regain US competitive advantages in the world cotton market.

As mentioned above, the US share of the cotton market has decreased over the last 10 years, and breeding tools need to be developed to combat this issue. Marker Assisted Selection (MAS) is a valuable tool that has thus far been underutilized. Overall MAS can result in more efficient and faster breeding – which accelerates the rate of improved cotton cultivars going to cotton growers. MAS can be best utilized in early generations because plants with undesirable genes can be eliminated early, thus reducing the time and space required to evaluate the remaining lines (Collard and Mackill, 2008). Before breeders can take full advantage of MAS, foundation research needs to be conducted such as the development and validation of high density genetic maps - constructed using multiple environmental replicated phenotyping experiments. This research uses such a method to develop a high density genetic map. The construction of

high density genetic maps via SNP markers along with the identification of QTLs that control major fiber quality and yield traits will be helpful in the implementation of MAS into breeding programs. Development and implementation of these tools into US cotton breeding programs will ultimately improve the United States competitiveness on the global cotton market.

**Objectives**

The objectives of this research project are to:

1. Grow and evaluate a previously genotyped and phenotyped RIL population at multiple locations in the U.S cotton belt.
2. Add phenotypic data from multiple environments to a pre-existing genetic map.
3. Identify QTLs related to yield traits such as lint percent and seed index.
4. Examine GxE interaction effects on yield traits such as lint percent and seed index.

# CHAPTER II

# LITERATURE REVIEW

## Types of DNA Markers

Many advancements have been made in the last 30 years regarding DNA molecular markers. These advancements have made it possible to identify differences among DNA sequences and resulted in genetic markers, which are associated with a specific loci that can be used in genetic mapping and plant breeding programs. These markers are based on DNA polymorphisms – differences in the nucleotide sequence such as single base pair changes, deletions, insertions, additions or patterns. DNA markers are easily detected, widely distributed throughout the genome, environmentally stable, and detectable at any stage of the plant growth and development. Characteristics that have made them heavily utilized in plant genetics. Some uses of molecular markers are genetic mapping, selection of parent breeding material, identification of species, and QTL analysis. Molecular markers can be used for Marker-Assisted Selection (MAS). MAS can result in more efficient and faster breeding, which accelerates the rate of improved cotton cultivars going to cotton growers.

It is important to note the original markers used in QTL mapping studies were AFLP, RFLP, RAPD, and SSRs. However, most of these markers have limited portability and utility in marker assisted breeding (Rong et al., 2004).

RFLP

RFLP (Restriction Fragment Length Polymorphism) was one of the earliest molecular markers used in genetic mapping. An RFLP is a specific sequence of DNA with a restriction site on each end and a target sequence in the middle. A probe sequence is used which binds to the target sequence by forming complementary base pairs. The probe is usually tagged with an enzyme or radioactively, so when it binds to the target sequence it can be easily identified/detected.

RFLPs produce a series of bands when a Southern blot (Southern, 1975) is performed with a particular combination of restriction enzyme and probe sequence. RFLP probes are often species-specific single locus probes and are about 0.5-3.0 kb long and can be obtained from a cDNA library or a genomic library.

RFLPs are codominant and locus specific markers, their coupling phase can be detected because the DNA fragments from all homologous chromosomes are detected. They can be reliably used in linkage analysis and can easily determine if a linked trait is present in an individual in a homozygous or heterozygous state. This information is highly desirable for recessive traits (Winter and Kahl, 1995). RFLPs have been successfully used in QTL mapping for a variety of crop species ranging from maize (*Zea mays*) (Burr et al., 1988; Murray et al., 1988) to peach (*Prunus persica*) (Eldredge et al., 1992). From 1988 to early 2000, RFLPs were widely used as the basic tool for genome mapping and other plant genetic research, specifically those crops with complex genomes or low levels of polymorphisms, such as cotton. Numerous studies have been conducted on the many ranges of cotton species using RFLPs, such as the origin and

evolution, population genetics, phylogenetic relationships, genome mapping, and QTL analysis (Wendel et al., 1989; Wendel and Albert, 1992; Meredith, 1992; Wang et al., 1992; Cantrell and Davis, 1993; Paterson et al., 1993; Wing, 1993; Kohel et al., 2001; Reinisch et al., 1994; Brubaker and Wendel, 1994; Shappley et al., 1996, 1998a, 1998b; Brubaker et al., 1999a, 1999b; Small and Wendel, 1999; Ulloa et al., 2000, 2005; Ulloa and Meredith, 2000; Jiang et al., 2000; Paterson et al., 2003; Rong et al., 2004, 2007).

While RFLPs were used in the past, several aspects of the marker make them difficult to use and the research community has abandoned them in favor of more efficient systems. RFLP require a large amount of high-quality DNA for restriction digestion and Southern blotting. They also require the presence of radioactive isotypes, which make research expensive and hazardous. The assay is time-consuming because of its low throughput, and is labor-intensive as the processes are difficult to automate. Often times only a low percent of markers are polymorphic, which limits their use especially for interspecific crosses. Their inability to detect single base changes restricts their use in modern plant breeding and makes them a poor candidate for use in MAS.

RAPD

RAPD (Random Amplified Polymorphic DNA) and was the first PCR-based technique (Williams et al., 1990). PCR (Polymerase Chain Reaction) is a technique that has been widely used in plant genetics and molecular biology since it was invented in 1983 (Bartlett and Stirling, 2003). RAPD markers do not require any specific knowledge of the DNA sequences of the source organism. Instead it uses a single DNA primer to directly amplify discrete random sequences depending on what parts of the source sequence are complementary to the primers. PCR is then used to amplify the DNA segments and they are run on a gel to separate them based on the fragments size. The fragments are visualized on the agarose gel by staining them with ethidium bromide.

RAPDs provide only dominant markers and it is not possible to determine if a DNA fragment is amplified from a locus that is heterozygous or homozygous (Williams et al., 1990). RAPD markers do not require prior sequence information or DNA probes to design specific primers, they also don't require the creation of genomic libraries, the time consuming southern blotting, or the use of radioactive isotopes, all of which are required when using RFLPs. RAPD markers are capable of detecting several loci with a single PCR cycle, and require small amounts of DNA (Williams et al., 1990). The procedure can be easily automated, and is capable of detecting higher levels of polymorphisms when compared with RFLP. This makes it useful for genetic studies for complex plants such as cotton. Previously RAPDs have been successfully used in the genetic mapping of several species including arabidopsis (*Arabidopsis thaliana*), peach (*Prunus persica*), lettuce (*Lactuca sativa*), and tomato (*Solanum lycopersicum*) (Reiter et

al., 1992; Chaparro et al., 1994; Kesseli et al., 1994; Grandillo and Tanksley, 1996). RAPDs have also been used with cotton for germplasm evaluations, identification of gene functions, genetic mapping, and QTL mapping (Multani and Lyon, 1995; Tatineni et al., 1996; Iqbal et al., 1997; Khan et al., 2000; Zhang et al., 2002). RAPDs have some limitations such as low reproducibility because of differences between DNA concentrations, PCR machines, and primer qualities. RAPDs can also have mismatches between the primer and the template DNA sequence which will result in no PCR product. As previously mentioned, RAPDs are also incapable of detecting allelic differences in heterozygotes.

AFLP

AFLP (Amplified Fragment Length Polymorphism) is a PCR-based marker technique that was developed in the early 1990's by a commercial company named Keygene. Keygene was and still is actively developing molecular and biotech tools for the plant breeding industry. The AFLP technique involves digesting DNA with two different restriction enzymes, followed by the ligation of adaptors to the sticky ends of the resulting fragments. The fragments are then amplified via PCR using complementary primers. Amplified fragments are visualized on polyacrylamide gels using either autoradiography or fluorescent labeling (Vos et al., 1995).

The AFLP procedure has been useful because it can generate a large number of mappable loci with a single amplification, which results in rapid genome coverage of markers. In comparison to RFLPs, AFLPs result in a higher marker density coverage in

the same population and even similar regions of the genome (Huang et al., 1994; Maheswaran et al., 1997). AFLP markers only show dominant relationships, similar to RAPDs (Meksem et al., 1995; Maughan et al., 1996). AFLPs are more reproducible and have better resolution across the genome compared to other methods. They do not require any prior sequence information for amplification (Mueller and Wolfenbarger, 1999; Meudt and Clarke, 2007). AFLP markers have been successfully used across a wide range of plants. Specifically in cotton, AFLPs have been used to estimate genetic diversity (Abdalla et al., 2001; Rana and Bhat, 2004; Zhang et al., 2005), develop linkage maps (Brubaker and Brown, 2003; Lacape et al., 2003; Mei et al., 2004; Zhang, et al., 2005), and conduct QTL analysis (Lacape et al., 2005). One of the limitations of AFLPs is they usually require polyacrylamide gels and the technique requires more time and labor than other methods. AFLPs were mostly used from 2000 to 2005 for mapping purposes because of their low cost, but were quickly abandoned as other more affordable molecular marker methods became available.

SSR

SSR (Simple Sequence Repeats), also known as microsatellites, are randomly tandem repeats of short nucleotide motifs, repeated several to over one hundred times. SSRs are widely distributed throughout the genomes of all eukaryotes. Microsatellite markers have been developed for many plant species, which allows SSR techniques to be easily applied for genetic mapping.

Primers are developed by cloning random segments of DNA from the target species. The library of clones is screened with fluorescently labeled oligonucleotide sequences, successful clones will be obtained if hybridization occurred between the oligonucleotide and a microsatellite repeat. After identifying potentially useful microsatellites, flanking sequences can be used to design primers which will amplify the SSR marker in a PCR reaction (Liu et al., 1996).

SSRs are reproducible, co-dominantly inherited, locus specific, widely distributed, relatively simple, and inexpensive. Because SSRs often reveal ample amount of polymorphisms, they tend to be more effective than other molecular marker systems in species with low genetic diversity. SSR markers have proven to be a valuable tool and used extensively in cotton genome mapping since the early 2000s. One of the first people to use SSR markers was Liu et al., (2000), who amplified 71 marker loci to a specific cotton genome using 65 SSR primer pairs. The protocol developed by Liu et al., (2000) and the SSR markers they developed helped to set the stage for further use of linkage mapping with additional SSR markers to expand the marker coverage of the cotton genome. Since then, many genetic maps and QTL analysis studies have been conducted using SSR markers (Mei et al., 2004; Lacape et al., 2005; Lin et al., 2005; Shen et al. 2005; Yu et al., 2012; Qin et al., 2015).Using SSRs, Guo et al., (2007) was able to map 1790 loci across all 26 cotton chromosomes with an average inter-loci distance of 1.91 cM.

While SSRs have been widely used, specifically for genetic mapping, they require nucleotide information to design primers, and the marker development is a labor-

intensive process. SSRs can be portable to closely related species, however, as genetic diversity increases, the percentage of successfully amplified loci goes down (Jarne and Lagoda, 1996). SSRs are still extensively used today; however, they are often used in conjunction with other markers such as AFLPs, RAPDs, and RFLPs in order to achieve good genome coverage (Lacape et al., 2003; Mei et al., 2004; Lin et al., 2008; Zhang et al., 2009). Tang et al., (2014) was able to construct a genetic map with 1540 SSR markers, but suggested that SSRs have limited use in the construction of high density genetic maps for cotton because the SSRs were unevenly mapped to chromosomes and poorly distributed.

SNP

A Single Nucleotide Polymorphism (SNP) marker relies on differences in a single nucleotide base between two DNA sequences. They are the most abundant sequence variation throughout most genomes. One SNP can be found every 100-300 base pairs (bp) in plants (Edwards et al., 2007). SNPs are useful to breeders because polymorphisms can be present in both coding regions and non-coding regions – in the proximity of nearly every gene. Unlike all other previously discussed molecular markers, SNPs do not require the use of agarose gels to run assays. This, combined with their stability and abundance through the cotton genome, make SNPs ideal candidates as molecular markers for use in marker-assisted selection (Van Deynze et al., 2009).

Development of high throughput sequencing and next generation technologies made the use of SNPs much more appealing to researchers. Companies such as Illumina

and 454 Life Sciences have developed sequencing platforms that are not only able to generate ten times the data in half the time, but also reduce the cost of discovery per SNP and cut down on labor because most of the sequencing is automated. These advancements have greatly reduced the price of sequencing and made using SNP markers much more affordable. SNP markers are co-dominant or dominant and PCR based, but do not require agarose gel to run assays. This gel free marker is well suited for marker assisted breeding with a large number of plants because breeders don't have to run a gel for every plant that is genotyped. Comparisons of SNPs and among the previously discussed DNA markers are summarized in Table 1.

SNP marker development has been used in many crops for genetic mapping. Some of the first crops to utilize SNP markers were rice (*Oryza sativa*) (Feltus et al., 2004), wheat (*Triticum aestivum*) (Somers et al., 2003; Zhang et al., 2003; Caldwell et al., 2004; Mochida et al., 2004), maize (*Zea mays*) (Ching et al., 2002; Batley et al., 2003), soybean (*Glycine max*) (Zhu et al., 2003; Kim et al., 2005), and barley (*Hordeum vulgare*) (Kanazin et al., 2002; Bundock and Henry, 2004). It wasn't until more recently that crops with more complex genomes such as cotton were able to capitalize on SNP markers. Several studies have been conducted and discovered over ten thousand SNPs in cotton; however, few SNPs have been able to be genetically mapped. Byers et al., (2012) was able to map 346 SNP markers based on an intraspecific F2 population and Yu et al., (2012) was able to map 247 markers based on a 186 interspecific RIL population. Gore et al., (2014) was able to map 491 SNP markers based on a 98 intraspecific RIL population. Hulse-Kemp et al., (2014) found 62,832 SNPs for five wild species that can

be used to help integrate new germplasm into cultivated G. *hirsutum*. Most recently, Wang et al., (2015) were able to map 4,049 SNP markers based on 59 F2 interspecific derived individuals and released the most comprehensive map to date, utilizing the more SNPs on a magnitude of 10x than any other research published. Of the 6442 SNPs mapped, few have been associated with cotton fiber traits.

**Table 1.** Comparison of common used DNA genetic markers. (Yu, 2009).

| FEATURE | RFLP | RAPD | AFLP | SSR | SNP |
|---|---|---|---|---|---|
| **Inheritance** | Co-dominant | Dominant | Dominant | Co-dominant | Co-dominant or dominant |
| **Pattern Detected** | Single-locus | Multi-Loci | Multi-Loci | Single-locus | Multi-Loci |
| **Cloning** | Required | No | No | Required | No |
| **Radioactivity** | Required | No | Required | No | No |
| **DNA quantity** | Large amount | Small | Moderate | Small | Small |
| **PCR-based** | No | Yes | Yes | Yes | Yes |
| **Sequence** | No | No | No | Required | Required |
| **Polymorphism** | High | High | Higher | Higher | Very High |
| **Ease of use** | Not easy | Easy | Easy | Easy | Easy |
| **Reproducibility** | High | Unreliable | Moderate | High | High |
| **Cost** | High | Low | Moderate | Low | Low |

## Types of Mapping Populations

There are several concerns that come to mind when developing a mapping population. The generation of the population (F2, Backcross, or RIL), population size - what is the right number of individuals in the population to create a thorough genetic map, and the type of cross used to develop the population: intra-specific (G. *hirsutum* x G. *hirsutum*) or an inter-specific (G. *hirsutum* x G. *barbadense*).

### Mapping Population Generation

The simplest form of a mapping population is an F2 population. Two parents are selected for a contrasting trait of interest and crossed to get a F1 population. Individual F1 plants are then selfed to produce the F2 population. Segregation ratios for codominant markers are 1:2:1 and segregation ratios for dominant markers are 3:1. The reasons that F2 populations are so widely used is because they can be developed with little effort and require less time to develop than other population types. This quick population development makes them well suited for preliminary mapping studies. However, in F2 populations, each plant represents one individual genotype as subsequent crossing or selfing cycles will cause segregation to continue to occur. Thus replications over different years, or different locations cannot be conducted. Since each plant represents one genotype, it is difficult to map quantitative traits. Certain traits like yield or fiber qualities cannot be reliably measured based off measurements from a single plant.

Another simple type of mapping population is a backcross population. Two parents are selected for their contrasting traits again and crossed to get a F1 population. The individual F1 hybrid plants are then crossed back to one of the parents, usually the parent with the recessive trait of interest. This method takes relatively little time to develop and the elite combination of the parents will still produce an elite genotype at the end of backcrossing. Some of the drawbacks of using backcross populations are similar to those of the F2 population. Quantitative traits cannot be mapped because one plant represents a single genotype and the population cannot be replicated over years and locations because the population will still be segregating for traits.

The last type of population used, which has gained more attention as of late, is a recombinant inbred line (RIL). A RIL population is developed by continuously selfing an individual member of an F2 population until it becomes completely homozygous. RILs have segregation ratios of 1:1. The biggest reason that RILs have gained in popularity for use in genetic mapping is because they are considered immortal – complete homozygosity has theoretically been achieved. The lines can be selfed and then planted across both years and locations and they will have the same genotype as the previous year (Yu et al., 2012). Statistically speaking, RIL populations are better than F2 populations to use for genetic mapping (Ferreira et al., 2006) because RILs are the product of multiple meiotic cycles whereas F2 populations are the product of only one meiotic cycle, multiple meiotic cycles allows more recombination events and better mapping resolution than the other populations. The reason that RIL lines didn't make a large appearance in plant genetic mapping sooner is that they require many seasons to

develop to ensure that lines are homozygous. The development of RIL lines can also be difficult in plants that have high inbreeding depression – which is minimal in cotton compared to other crops such as maize or sorghum (*Sorghum bicolor*).

## Population Size

Choosing the population size when developing a mapping population is important. A population needs to be large enough to ensure better QTL detection power but small enough that phenotyping and genotyping can be accomplished quickly and efficiently. As the pressure for more precise genetic maps and higher levels of confidence grows, researchers have started using larger populations. Using a RIL population can help cut down on the required population size because more recombination has occurred. This results in better map resolution and also allows the mean of a phenotype to be measured across environments since the genotypes are homozygous. Other important factors to consider are the heritability of the trait, and the magnitude of the QTL effect that is being identified. The problem of using a small population could result in failure to identify QTLs and exaggerate the effect of the identified QTLs (Beavis, 1998). It is suggested that a population size smaller than 300 individual plants may preclude the estimation of the distribution of QTL effects (Erickson et al., 2004). However, results from (Ferreira, 2006) indicate that a population size of 200 is sufficient to construct an accurate genetic map.

Intra-specific vs Inter-specific

There are two main types of hybrid populations used in QTL mapping – intraspecific populations, obtained from crossing two G. *hirsutum* lines together, and interspecific populations, obtained from crosses between G. *hirsutum* and G. *barbadense*. Intraspecific populations were originally used early on for mapping studies because upland cotton lines represent the major source of breeding material in the US (Shappley et al., 1998). However, intraspecific populations are difficult to use because there is a low number of polymorphisms among upland cotton cultivars and thus the maps have a low density of marker coverage.

More recently, interspecific crosses between G. *hirsutum* and G. *barbadense* have been used to develop genetic maps because of their high DNA polymorphisms and broad genetic variation in fiber traits (Lin et al., 2005). To help make interspecific maps more applicable and useful for commercial cotton breeding, several studies have utilized restriction site-associated DNA sequencing to genotype commonly used commercial cotton lines (Peterson et al., 2012). This allows SNPs mapped from the mapping population to be easily transferred to the commercial cotton breeding lines (Byers et al., 2012).

**QTLs**

A quantitative trait locus (QTL) is a section of the DNA strand which makes up a gene, which correlates with a difference in the phenotypes of quantitative traits such as yield or lint percent. What makes QTLs useful in breeding is if the gene(s)/QTL that

control a trait can be identified and located in the genome, then hypothetically the gene can be cloned and inserted into desired breeding material. Another significant use for QTLs, and one of the biggest focuses in recent plant breeding research, is marker assisted selection (MAS). If the gene(s) controlling a trait can be identified along with molecular markers that are in close proximity to the gene, then large populations can be screened for the presence of the molecular markers and it can be determined if the gene of interest is present or not. This method allows researchers and breeders to screen large populations for the presence of the gene without growing the plants full to maturity – which results in faster selections and fewer resources put into growing undesirable plants.

The focus of identifying and utilizing quantitative trait loci (QTLs) to be used in marker assisted selection has been the goal for most breeding programs. The first QTLs to be associated with fiber traits were identified by Shappley et al., (1998). Since then, many QTL mapping studies have been conducted and thousands of QTLs have been identified. Generally, QTLs should be mapped to intervals of < 5.0 cM and if possible use flanking markers to greatly improve reliability. If the mapping intervals are greater than 5.0 cM marker assisted selection can be ineffective. There are many methods of identifying QTLs with the simplest being an ANOVA and the more complex being software such as IciMapping and QTLCartographer.

Previous Genetic Maps

The first genetic map to be constructed for cotton was done by Reinisch et al. (1994), using an interspecific cross. However, Shappley et al., (1996) was the first to create a genetic map using an intraspecific cross. Since then, dozens of genetic maps have been created for both intraspecific crosses and interspecific crosses (Zhang et al., 2002; Lacape et al., 2003, 2009; Rong et al., 2004; Blenda et al., 2006; Frelichowski et al., 2006; Guo et al., 2007; Yu et al., 2007, 2008, 2011, 2012, 2013; Lin et al., 2008; Chen et al., 2009; Van Dyanzye et al., 2009; Zhang et al., 2009; Xiao et al., 2009; Byers et al., 2012; Tang et al., 2014; Li et al., 2015; Wang et al., 2015), these maps collectively consist of up to 3,500 loci and 2,500 DNA markers, spanning ~5,500 cM with an average interval of about 1.3 cM or 573 kb between loci. In addition, several reference maps have been constructed, including 'TM-1' (Lee et al., 2013; Zhang et al., 2012) and 'Acala Maxxa' (Tomkins et al., 2001). Reference maps are used in conjunction with anchor markers to align the sequence reads in the correct orientation. This allows a sense of congruency between maps, allows them to be compared to one another, and validates that the researcher has constructed the new map correctly in comparison to the reference map.

'Cottongen.org' currently has a database consisting of 50 publicly available genetic maps which is a valuable tool when constructing a genetic map. The current genetic maps still leave many questions unanswered, which is why researchers are focusing on constructing more complete and comprehensive maps. The problems with most of the previous maps is that most of the QTLs were mapped with RFLP, RAPD and

AFLP (Jiang et al., 1998; Kohel et al., 2001; Paterson et al., 2003; Mei et al., 2004; Chee et al., 2005a, b; Draye et al., 2005; Frelichowski et al., 2006; Zhang et al., 2009). These DNA markers are not well suited for marker-assisted breeding. Furthermore, most of the QTLs were mapped to intervals of > 5.0 cM, so the flanking DNA markers were not close enough for effective marker-assisted selection (Jiang et al., 1998; Kohel et al., 2001; Paterson et al., 2003; Mei et al., 2004; Chee et al., 2005a, b; Draye et al., 2005; Frelichowski et al., 2006; Zhang et al., 2009). Additionally, most used early-generation segregating populations derived from interspecific crosses between *G. hirsutum* and *G. barbadense* due to the lack of immortal RIL populations at that time and were mapped using low DNA marker maps. Finally, most of the cotton genetic maps were constructed and QTLs were mapped using interspecific *G. hirsutum* x *G. barbadense* populations while most national cotton breeding programs are practiced within a species (*G. hirsutum* x *G. hirsutum*). This also limits the utility of the existing DNA markers in cotton breeding. Therefore, it is necessary to advance cotton fiber QTL research such as the development of co-segregating or fiber gene-specific SNP markers and also to translate the marker interspecific polymorphism information into intraspecific polymorphism information so that the markers could be effectively used in cotton fiber breeding programs.

The assumption is that a better genetic map will result in the better implementation of MAS, which in-turn results in meaningful phenotypic differences observed in earlier generations during the selection process and an overall improvement in how plant breeders develop new material. However, not all genetic maps are created

equal: the number of markers, marker spacing, and the size of the population used to develop the genetic map play a role in how useful the genetic map will be. The number of markers used to create a genetic map is important to its accuracy. The more markers that are on a genetic map, the more likely it is that one of the markers will be located close to a gene of interest – this allows researchers to pinpoint the location of the gene. If there are not enough markers, the likelihood of accurately identifying a QTL/gene is reduced. Marker spacing is also important as evenly spaced markers will help zero in on the position of the QTL. A QTL can be identified between markers, but if the distance between the markers is large, then the accuracy of the map decreases because is difficult to accurately pinpoint the exact location of the QTL. An abundance of evenly distributed markers is preferable because it not only gives high density coverage, but also has even spacing across all chromosomes. The effect of population size on the construction of genetic maps is widely recognized. In some cases, too small of a population results in zero QTLs being identified (Rami et al., 1998). Small population size can lead to underestimation of QTL number, and overestimation of the QTL effects (Beavis, 1998). Beavis also reported that using a small number of lines for QTL analysis does not accurately identify the location or magnitude of the QTL effect, especially if there are a large number of small effect QTLs segregating in the genome. Increasing the population size results in an increase in the cost of QTL analysis. Smaller populations historically were used because the cost to genotype populations was high; since that is no longer the case due to advancements in technology which have dropped the price of genotyping by a rate of 10x, researchers can now phenotype large populations.

Despite the number of studies on this topic, most maps still do not have an adequate marker density to support marker-assisted selection breeding or construct high-resolution maps and few SNPs have been associated with cotton fiber traits (An et al., 2007, 2008, 2009; Lu et al., 2009; Xiao et al., 2009; Fang et al., 2010; Byers et al., 2012; Yu et al., 2012).

QTLs Identified

QTLs for a small number of fiber yield and quality component traits have been mapped in cotton (Jiang et al., 1998; Kohel et al., 2001; Paterson et al., 2003; Mei et al., 2004; Chee et al., 2005a, b; Draye et al., 2005; Frelichowski et al., 2006; Zhang et al., 2009). Additionally, 'cottongen.org' has a database that includes 988 QTL loci of 200 QTL trait data that they have collected from publications and plan to release 1,000 further QTLs in the near future. However, as previously described, most maps used to identify QTLs had low marker numbers, used early generation segregating populations from interspecific crosses, and failed to use immortal RIL populations. Results of these efforts have provided useful information about the positions of the genes controlling fiber traits. However, when Draye et al., (2005) compared the mapping results for fiber micronaire and fineness with those of Paterson et al., (2003), they found that only ~1/3 of the QTLs (41 by Draye et al., 2005 and 25 by Paterson et al., 2003) were consistent. The discrepancy between Draye and Paterson experiments could be the result of QTLs which interact with environmental factors such as locations and growing seasons. Therefore, it is necessary to validate and refine the QTLs using a permanent mapping

population and high-density markers that allow phenotyping the traits in multiple environments.

## Future of QTLs

Despite the many advances in marker development, QTL identification, and population development, there is still much work that needs to be done before QTLs can be effectively used in commercial cotton breeding. There are still some hurdles that need to be overcome before successful implementation of QTLs and MAS such as a reduction of the cost of genotyping. While the cost of genotyping per data point has dropped substantially in the last ten years, the price per sample is still rather high. If one considers a SNP chip with 70,000 SNPs and a cost of $65.00 per sample, the cost per data point is less than $0.01. However, in large commercial breeding programs there are hundreds of F2 plants being evaluated which can lead to costs exceeding $32,500.00 per population. The price combined with the cost of DNA extraction, and phenotyping can exceed $40,000.00 per population. This cost needs to decrease significantly before MAS can be reasonably used in cotton breeding programs.

The other hurdle is that the location and effect of QTLs need to be confirmed and validated in additional genetic backgrounds before it can be used across breeding programs. This requires more precise maps and the ability to generate large amounts of molecular data to construct those maps (Varshney et al., 2009).

# CHAPTER III

# MATERIALS AND METHODS

## Breeding Material

The recombinant inbred line (RIL) population at the F3:8 generation was developed by the Cotton Improvement Laboratory, Department of Soil and Crop Sciences, Texas A&M AgriLife Research, College Station, Texas. The population was developed from the interspecific cross between 'TAM 94L-25' (Smith, 2003) (G. *hirsutum*) x 'NMSI 1331' (Roberts et al., 1997) (G. *barbadense*). These parents were chosen because of their high DNA polymorphism and broad genetic variation for fiber traits such as: lint yield, lint percent, seed weight, fiber length, strength, micronaire, length uniformity, and elongation. In 2009, 2010, and 2011 the entire populations consisting of 198 RILs and both parents were grown in a randomized complete block design (RCBD) with a non-replicated trial in 2009 and three replicates in 2010 and 2011 at the Texas A&M AgriLife Research Farm near College Station, Texas (Liu, 2014). In 2012, 2013, and 2014 176 RILs and both parents were grown for seed increase at the Texas A&M AgriLife Research Farm, College Station, Texas, to replenish seed inventory. In 2015, 176 RILs, both parents, and two check cultivars were grown in a randomized complete block design (RCBD) with two replicates at three locations: Texas A&M AgriLife Research Farm, College Station, Texas; Texas A&M AgriLife Research Farm, Lubbock, Texas; Louisiana State University AgCenter Northeast Research Station, St. Joseph, Louisiana. Standard cotton production practices were used to grow these field trials. The soil at Texas A&M AgriLife Research Farm in College Station,

Texas is a Westwood silt loam. The soils of the Westwood series are coarse-silty, mixed, thermic Udifluventic Haplustepts. The soil at Texas A&M AgriLife Research Farm in Lubbock, Texas is a Pullman clay loam. The soils of the Pullman series are fine, mixed, superactive, thermic Torrertic Paleustolls. The soil at Louisiana State University AgCenter Northeast Research Station in St. Joseph, Louisiana is a Commerce silt loam. The soils of the Commerce series are fine-silty, mixed, superactive, nonacid, thermic Fluvaquentic Endoaquepts. 2015 was an uncharacteristically wet year, with rainfall totals exceeding the annual averages (Table 2).

**Table 2.** 2015 monthly rainfall totals and historical averages for the three planting locations: College Station and Lubbock, Texas, and St. Joseph, Louisiana (US Climate Data, 2016).

| Location | College Station, Texas | | Lubbock, Texas | | St. Joseph, Louisiana | |
|---|---|---|---|---|---|---|
| Month | 2015 monthly rainfall (in) | Average rainfall (in) | 2015 monthly rainfall (in) | Average rainfall (in) | 2015 monthly rainfall (in) | Average rainfall (in) |
| January | 6.66 | 3.23 | 1.34 | 0.71 | 6.21 | 5.47 |
| February | 0.75 | 2.83 | 0.77 | 0.63 | 3.40 | 5.35 |
| March | 5.83 | 3.19 | 0.60 | 1.14 | 8.56 | 5.55 |
| April | 4.80 | 2.68 | 3.01 | 1.65 | 4.13 | 4.41 |
| May | 9.73 | 4.33 | 13.24 | 2.80 | 6.56 | 4.76 |
| June | 5.21 | 4.45 | 3.42 | 3.19 | 3.08 | 4.13 |
| July | 0.31 | 2.13 | 3.81 | 2.40 | 2.26 | 4.41 |
| August | 1.36 | 2.68 | 1.51 | 2.24 | 0.37 | 3.98 |
| September | 1.74 | 3.19 | 0.36 | 2.17 | 1.95 | 3.50 |
| October | 8.81 | 4.92 | 4.60 | 1.73 | 9.00 | 3.98 |
| November | 5.03 | 3.23 | 1.13 | 0.91 | 9.94 | 5.47 |
| December | 8.01 | 3.23 | 9.92 | 0.79 | 1.63 | 5.43 |
| Total | 58.24 | 40.09 | 43.71 | 20.36 | 57.09 | 56.44 |

**Phenotypic Data Collection**

When matured, 30-50 bolls were hand-harvested from each plot and ginned at the Cotton Improvement Laboratory, Texas A&M AgriLife Research, College Station, Texas. Before and after ginning, each sample was weighed to determine lint percent. After ginning, 100 fuzzy seeds were weighed on an electronic scale to determine seed index. Additionally, 50 grams of cotton fiber from each sample was sent to the Fiber and Biopolymer Institute at Texas Tech University where fibers were measured with HVI. From the HVI measures, fiber properties were obtained, including: upper half mean length (UHML), micronaire, strength, length uniformity, and elongation. However, this study will only focus on yield traits: lint percent and seed index. The HVI fiber quality traits will be examined in an alternate study.

**Genotypic Data Collection**

Plant tissue was collected in 2009 by Yun-Hua Liu, a previous graduate student with the Cotton Improvement Laboratory. All sequencing was and development of the linkage map framework was done by Dr. Hongbin Zhang's lab at Texas A&M University. The BamH I sites of the RIL population and an additional 150 widely used U.S. breeding germplasm lines were sequenced by the RAD-seq technology following the procedures described by Peterson et al. (2012) with minor modifications. The sequencing data was analyzed using the pipeline developed by Peterson et al. (2012). The upland cotton genome sequence TM-1 was used as a reference genome sequence for the analysis courtesy of Monsanto (Yang et al., 2015). The upland cotton physical map

has already been developed from BIBACs constructed with BamH I and the present genetic map was integrated to the physical map using the physical map BAC end sequences as anchors (Zhang, 2015).

**Statistical Analysis**

Analyses of variance for lint percent and seed index were performed using SAS 9.4 (SAS Institute Inc., 2013). Each location and year was classified as an environment (Table 3).

**Table 3.** Classification of environments based on location and year components.

| Location | Year | Environment |
|---|---|---|
| College Station, Texas | 2009 | 1 |
| College Station, Texas | 2010 | 2 |
| College Station, Texas | 2011 | 3 |
| St. Joseph, Louisiana | 2015 | 4 |
| College Station, Texas | 2015 | 5 |
| Lubbock, Texas | 2015 | 6 |

Proc GLM with Environments and Genotypes as fixed effects and Reps as random effects, was used to determine differences among genotypes and between environments. Proc GLM procedure found significant differences between genotypes and environments for both lint percent and seed index (Tables 4-7), suggesting that a potential QTL exists for both traits. It also found that environments were highly significant suggesting that the variability of the environments may affect the ability to identify a QTL across all environments.

**Table 4.** ANOVA of lint percent distributions at St. Joseph, LA, and College Station, and Lubbock, TX environments from 2009-2015.

**ANOVA**

| **Source** | **df** | **SS** | **MS** | **F** |
|---|---|---|---|---|
| Rep | 2 | 0.009 | 0.0047 | 1.73 |
| Genotype | 175 | 0.807 | 0.0046 | 8.09** |
| Rep x Genotype | 343 | 0.218 | 0.0006 | 1.60** |
| Environment | 5 | 1.203 | 0.2407 | 75.74** |
| Environment x Rep | 5 | 0.014 | 0.0029 | 7.46** |
| Genotype x Environment | 838 | 1.084 | 0.0012 | 3.26** |
| Error | 678 | 0.269 | 0.0004 | |
| Total | 2048 | 4.470 | | |

*Significant at the 0.05 probability level

**Significant at the 0.01 probability level

**Table 5**. Environment averages of lint percent at St. Joseph, LA, and College Station, and Lubbock, TX environments from 2009-2015.

| **Environment** | **N** | **Means** | |
|---|---|---|---|
| Lubbock 2015 | 349 | 39.2 | a |
| College Station 2015 | 352 | 34.7 | b |
| College Station 2010 | 329 | 34.2 | c |
| College Station 2009 | 173 | 34.0 | c |
| Louisiana 2015 | 338 | 33.5 | d |
| College Station 2011 | 503 | 31.4 | e |

**Table 6.** ANOVA of seed index distributions at St. Joseph, LA, and College Station, and Lubbock, TX environments in 2015.

**ANOVA**

| **Source** | **Df** | **SS** | **MS** | **F** |
|---|---|---|---|---|
| Rep | 1 | 0.553 | 0.5537 | 0.22 |
| Genotype | 175 | 724.737 | 4.1413 | 9.32** |
| Rep x Genotype | 175 | 77.769 | 0.4443 | 1.02 |
| Environment | 2 | 636.160 | 318.0804 | 126.50** |
| Environment x Rep | 2 | 5.018 | 2.5093 | 5.78** |
| Genotype x Environment | 349 | 247.394 | 0.7088 | 1.63** |
| Error | 335 | 145.444 | 0.4341 | |
| Total | 1039 | 1884.257 | | |

*Significant at the 0.05 probability level

**Significant at the 0.01 probability level

**Table 7.** Environment averages of seed index at St. Joseph, LA, and College Station, and Lubbock, TX environments in 2015.

| **Environment** | **N** | **Means** | | | |
|---|---|---|---|---|---|
| Louisiana 2015 | 338 | 12.61 | a | | |
| Lubbock, TX, 2015 | 350 | 11.01 | | b | |
| College Station, TX, 2015 | 352 | 10.81 | | | c |

**Map Construction and QTL Analysis**

Linkage Map construction and QTL analysis was performed using QTL IciMapping 4.0 software (Meng et al., 2015). Linkage map framework, anchoring of markers, and construction of bins was done by Yun-Hua Liu, a previous graduate student at the Cotton Improvement Lab (Liu, 2014). No statistical difference was identified between reps within each environment, so for mapping purposes the trait data from the reps of each environment were averaged together. IciMapping 4.0 was employed to detect QTL for phenotypic variations of lint percent and seed index using inclusive composite interval mapping. A large number of SNP markers were identified and the BIN function was used to bin redundant markers, the MAP function was used to construct linkage maps of the biparental RIL population, BIP function was used to map additive, dominant, and epistatic genes, and the MET function was used to conduct a QTL-by-environment interaction analysis. Several parameters were input into the mapping software to describe the data for map construction and QTL identification. Within IciMapping, the population was identified as a RIL population, the mapping function was set at Kosambi, the marker space type was set to positions on the chromosome as opposed to intervals between markers, and the marker spacing unit was set to centi-Morgans (cM). Rstudio was used to create phenotypic distribution figures.

# CHAPTER IV

# RESULTS AND DISCUSSION

## Lint Percent Phenotypic Variation of the RIL Population

The lint percent of the TAM 94L-25 x NMSI 1331 RIL population was phenotyped in field trials over four years (Table 8). In 2009, the population had an average lint percent of 34.0%, ranging from 18.0 to 47.0%. In 2010, the field trial had three replicates and the population had an average lint percent of 34.1%, ranging from 18.0 to 46.0%. In 2011, the field trial also had three replicates and the population had an average lint percent of 31.4%, ranging from 18.3 to 41.6%. In 2015, the St. Joseph, Louisiana field trial had two replicates and the population had an average lint percent of 33.5%, ranging from 24.4 to 39.6%. In 2015, the College Station, Texas field trial had two replicates and the population had an average lint percent of 34.7%, ranging from 29.9 to 42.5%. In 2015, the Lubbock, Texas field trial had two replicates and the population had an average lint percent of 39.2%, ranging from 33.7 to 43.9% (Figures 1-7).

**Table 8.** Fiber lint percent (%) distributions of the RIL population during 2009-2015 at the St. Joseph, LA, and College Station, and Lubbock, TX environments.

| | 2009 CS | 2010 CS | 2011 CS | 2015 LA | 2015 CS | 2015 LU | Average |
|---|---|---|---|---|---|---|---|
| Sample Size | 173 | 144 | 176 | 175 | 176 | 176 | 170 |
| Minimum | 18.00 | 18.00 | 18.33 | 24.46 | 29.92 | 33.77 | 23.75 |
| Mean | 34.06 | 34.15 | 31.46 | 33.58 | 34.78 | 39.27 | 34.55 |
| Maximum | 47.00 | 46.00 | 41.67 | 39.68 | 42.56 | 43.90 | 43.47 |
| Range | 29.00 | 28.00 | 23.34 | 15.22 | 12.64 | 10.13 | 19.72 |
| Variance | 0.23 | 0.27 | 0.16 | 0.06 | 0.04 | 0.04 | 0.13 |
| StdError | 4.84 | 5.22 | 4.01 | 2.36 | 2.04 | 2.10 | 3.43 |
| P-value | 0.82 | 0.35 | 0.30 | 0.18 | 0.11 | 0.02 | 0.29 |

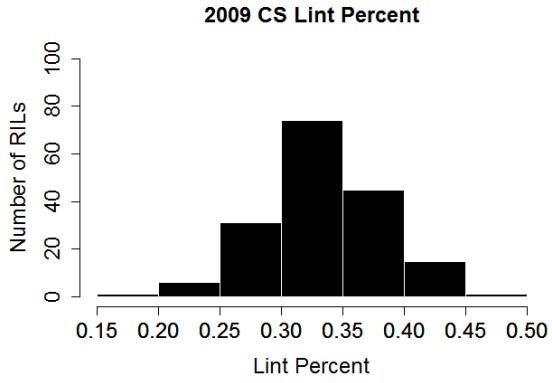


**Figure 1.** Lint percent distributions in the TAM 94L-25 x NMSI 1331 RIL population in 2009 at the College Station, Texas environment.

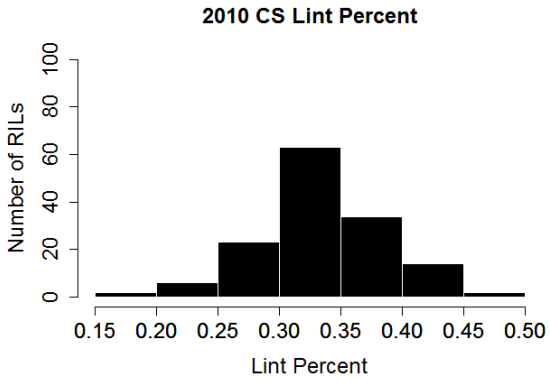


**Figure 2.** Lint percent distributions in the TAM 94L-25 x NMSI 1331 RIL population in 2010 at the College Station, Texas environment.

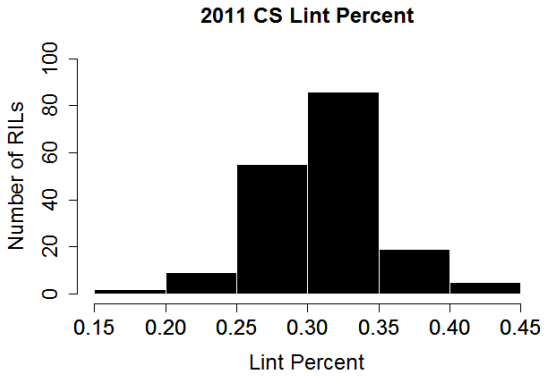


**Figure 3.** Lint percent distributions in the TAM 94L-25 x NMSI 1331 RIL population in 2011 at the College Station, Texas environment.

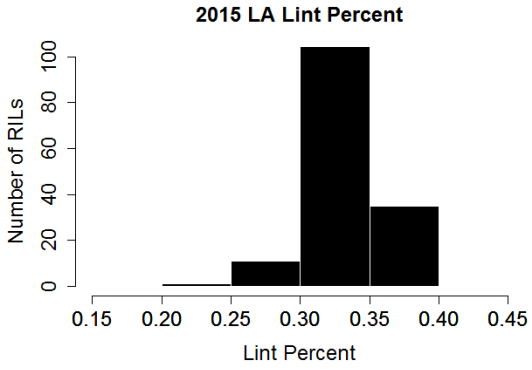


**Figure 4.** Lint percent distributions in the TAM 94L-25 x NMSI 1331 RIL population in 2015 at the St. Joseph, Louisiana environment.

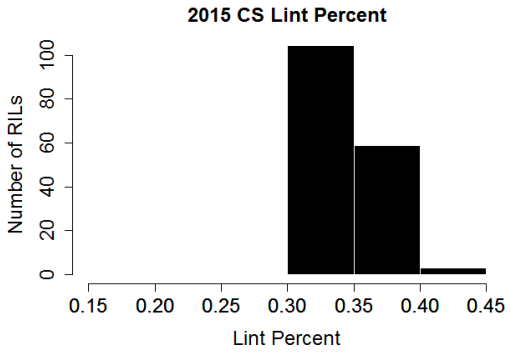


**Figure 5.** Lint percent distributions in the TAM 94L-25 x NMSI 1331 RIL population in 2015 at the College Station, Texas environment.

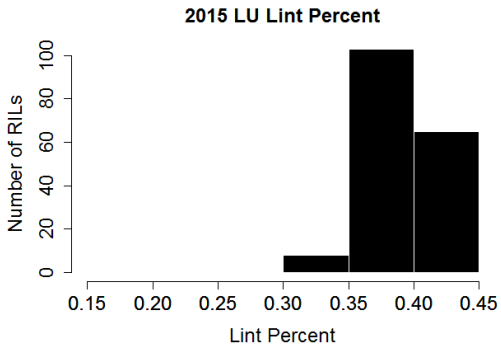


**Figure 6.** Lint percent distributions in the TAM 94L-25 x NMSI 1331 RIL population in 2015 at the Lubbock, Texas environment.

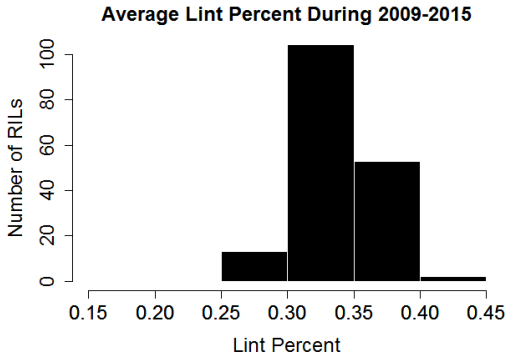


**Figure 7.** Average Lint percent distributions in the TAM 94L-25 x NMSI 1331 RIL population during 2009 – 2015 across all environments.

**Seed Index Phenotypic Variation of the RIL Population**

The seed index of the TAM 94L-25 x NMSI 1331 RIL population was phenotyped in field trials over one year at three locations (Table 9). In 2015, the St. Joseph, Louisiana field trial had two replicates and the population had an average seed index of 12.62, ranging from 9.90 to 15.78. In 2015, the College Station, Texas field trial had two replicates and the population had an average seed index of 10.81, ranging from 8.62 to 13.16. In 2015, the Lubbock, Texas field trial had two replicates and the population had an average seed index of 11.01, ranging from 8.49 to 14.06 (Figures 8-11).

**Table 9.** Seed index distributions of the RIL population during 2015 at the St. Joseph, LA, and College Station, and Lubbock, TX environments.

| | 2015 LA | 2015 CS | 2015 LU | Average |
|---|---|---|---|---|
| Sample Size | 175 | 176 | 176 | 176 |
| Minimum | 9.91 | 8.63 | 8.50 | 9.01 |
| Mean | 12.62 | 10.81 | 11.01 | 11.48 |
| Maximum | 15.78 | 13.17 | 14.06 | 14.34 |
| Range | 5.88 | 4.54 | 5.57 | 5.33 |
| Variance | 1.33 | 0.65 | 0.93 | 0.97 |
| StdError | 1.15 | 0.81 | 0.97 | 0.98 |
| P-value | 0.06 | 0.86 | 0.54 | 0.49 |

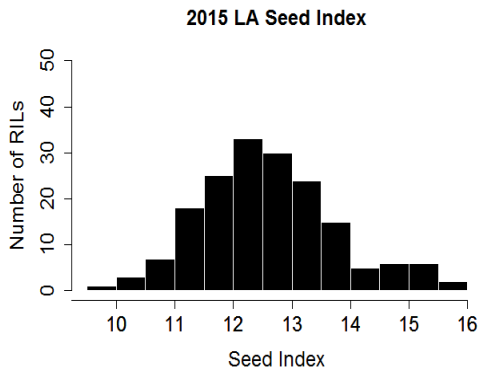


**Figure 8.** Seed index distributions in the TAM 94L-25 x NMSI 1331 RIL population in 2015 at the St. Joseph, Louisiana environment.

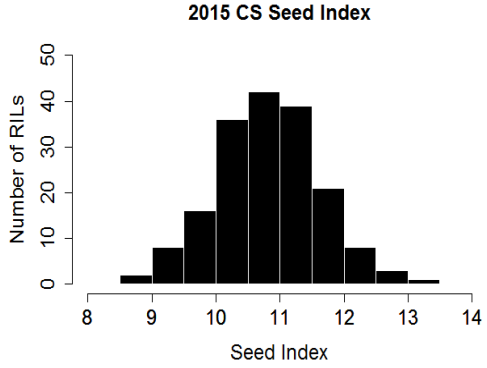


**Figure 9.** Seed index distributions in the TAM 94L-25 x NMSI 1331 RIL population in 2015 at the College Station, Texas environment.

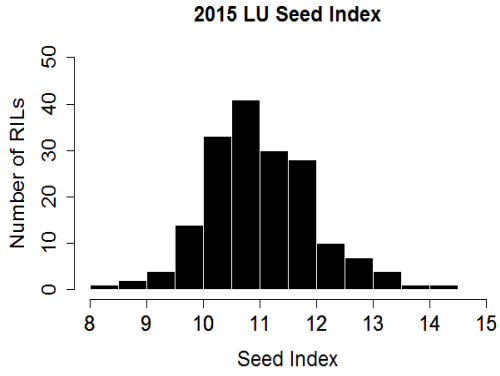


**Figure 10.** Seed index distributions in the TAM 94L-25 x NMSI 1331 RIL population in 2015 at the Lubbock, Texas environment.

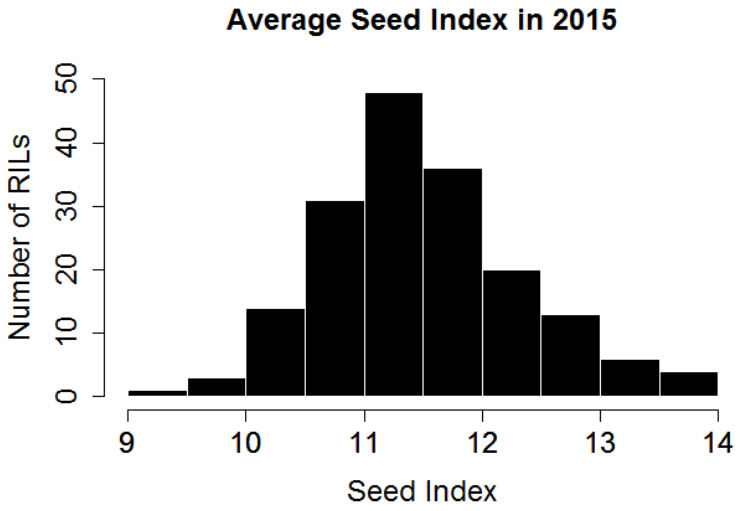


**Figure 11.** Average seed index distributions in the TAM 94L-25 x NMSI 1331 RIL population in 2015 across all environments.

**Mapping Results**

The RIL population was genotyped by ddRADseq, and nearly one million SNPs were identified. Due to conventional mapping software being unable to handle such a large number of SNPs, the SNPs were constructed into 1678 bins. An ultra-high-density SNP genetic map was developed for the cultivated tetraploid cotton. The 1678 bins are distributed among the 26 chromosomes of the tetraploid cotton, covering a total of 5054.2 cM, ranging from 124.24 cM to 276.74 cM per chromosome.

## QTLs Identified

Based on inclusive composite interval mapping (ICIM) method, a total of 18 potential QTLs were detected on 15 chromosomes, explaining 6.15-37.48% of the phenotypic variation (Tables 8-9), with LOD scores ranging from 2.50 to 6.22. In six environments, 14 QTLs for lint percent were identified, with one of the QTLs being detected in more than one environment. In three environments, 4 QTLs for seed index were identified.

## Lint Percent

A total of 14 QTL for lint percent were identified on 13 chromosomes (Chr1, Chr4, Chr5, Chr7, Chr9, Chr10, Chr13, Chr14, Chr18, Chr19, Chr24, Chr25, Chr26), explaining 6.30-11.91% of the PV, with LOD scores ranging from 2.50 to 3.96. Eleven QTLs showed positive additive effects, while four QTLs showed negative additive effects (Table 10).

**Table 10.** QTLs identified for lint percent in the RIL population, obtained from IciMapping 4.0.

| QTL | Chromosome | Position (cM) | LOD | PVE(%) | Additive | Environment |
|---|---|---|---|---|---|---|
| qLP-c1-1 | Chr1 | 75.0 | 2.76 | 7.08 | -0.024 | 2 |
| qLP-c4-1 | Chr4 | 0.0 | 3.96 | 8.07 | 0.012 | 3 |
| qLP-c5-1 | Chr5 | 40.0 | 3.10 | 8.52 | -0.045 | 2 |
| qLP-c7-1 | Chr7 | 248.0 | 3.36 | 11.91 | 0.015 | 6 |
| qLP-c9-1 | Chr9 | 30.0 | 2.50 | 7.02 | -0.023 | 1 |
| qLP-c10-1 | Chr10 | 165.0 | 2.77 | 7.62 | 0.016 | 6 |
| qLP-c13-1 | Chr13 | 219.0 | 3.06 | 8.47 | 0.014 | 1 |
| qLP-c14-1 | Chr14 | 96.0 | 3.03 | 7.28 | 0.022 | 1 |
|  | Chr14 | 96.0 | 2.88 | 7.72 | 0.023 | 2 |
| qLP-c18-1 | Chr18 | 29.0 | 2.57 | 6.30 | -0.022 | 1 |
| qLP-c19-1 | Chr19 | 76.0 | 3.12 | 8.28 | 0.023 | 1 |
| qLP-c24-1 | Chr24 | 37.0 | 2.86 | 7.15 | 0.009 | 5 |
| qLP-c25-1 | Chr25 | 54.0 | 3.17 | 7.68 | 0.025 | 3 |
| qLP-c26-1 | Chr26 | 42.0 | 3.68 | 8.91 | 0.026 | 3 |
| qLP-c26-2 | Chr26 | 163.0 | 2.99 | 6.30 | 0.017 | 3 |

Seed Index

A total of 4 QTL for seed index were identified on 4 chromosomes (Chr4, Chr8, Chr17, Chr19), explaining 6.15-37.48% of the PV, with LOD scores ranging from 2.51 to 6.23 (Table 11).

**Table 11.** QTLs identified for seed index in the RIL population, obtained from IciMapping 4.0.

| QTL | Chromosome | Position (cM) | LOD | PVE(%) | Additive | Environment |
|---|---|---|---|---|---|---|
| qSI-c4-1 | Chr4 | 62.0 | 2.51 | 37.48 | 0.987 | 4 |
| qSI-c8-1 | Chr8 | 133.0 | 2.93 | 6.15 | 0.333 | 5 |
| qSI-c17-1 | Chr17 | 26.0 | 6.23 | 13.59 | -0.302 | 5 |
| qSI-c19-1 | Chr19 | 21.0 | 3.60 | 8.85 | 0.420 | 5 |

Discussion

In the present study, a high-density genetic map comprising 1,678 loci was constructed covering 5,054.2 cM. This new map is currently longer than any other previously published inter-specific map including Rong et al., (2004) 4,447.9 cM, Yu et al., (2011) 3,380 cM, Yu et al., (2012) 4,418.9 cM, Zhao et al., (2012) 3,667.62 cM, and Wang et al., (2015) 4,042 cM.

This map is the product of one of the first applications of GBS on an inter-specific RIL population, thus it is not possible to compare the QTL and mapping results to those of other studies. However, with GBS being rapidly accepted, it will not be long until other studies are published and the resulting QTLs can be validated and compared.

Only one potential QTL for lint percent could be identified in multiple environments, suggesting a large GxE effect on these traits and the expression of related QTLs; this coincides with the fact that most environments were significantly different from each other. Many studies have reported difficulties of obtaining stable QTLs in multiple environments (Shen et al., 2006; Zhang et al., 2009; Zhang et al., 2012; Tang et

al., 2014). The environmental differences, specifically the drought in 2011 and the unusually wet 2015 season might have seriously affected yield performance in cotton.

The inability to detect the same QTL in every environment was expected, especially for quantitative traits like lint percent and seed index. Quantitative traits are complex and affected by environmental differences; the genes affecting lint percent in Lubbock, Texas are likely not the same genes controlling lint percent in St. Joseph, Louisiana.

The study by Wang et al., (2015) suggests that fiber quality traits are more stable than fiber yield traits in multiple environments, which is in agreement with the findings in the present study. It also provides encouragement that several stable QTLs for fiber quality traits might be identified for this same population once the data is analyzed.

With such extensive genome coverage and large population size, it is possible that there is enough statistical power to repeatedly identify QTLs with small effects; however, this is not the case with large effect QTLs, which are affected by environmental differences and gene-gene interactions.

The construction of larger mapping populations along with the development of higher coverage linkage maps is needed. Regardless, the identification of 18 potential QTLs related to lint percent and seed index greatly assists and propels the efforts to fully utilize MAS in commercial cotton breeding. It is most likely that because of the low explained phenotypic variation and the lack of stable QTLs identified across multiple years and environments that the QTLs identified in the present study are not related to

large effect candidate genes. The true usefulness and utility of these identified QTLs will need to be evaluated through further testing and application in MAS studies using multiple RIL populations.

The application and implementation of MAS has limitations, specifically due to the fact that GxE effects are severe and environments are vastly different. The same QTL is not detected in every environment, which means MAS may be able to be used when breeding for a specific target area with a consistent environment. However, MAS loses its utility when applied to large areas with different environments. It is unlikely that MAS could be successfully applied to cotton breeding across the state of Texas and Louisiana because the environments are drastically different.

# CHAPTER V

# CONCLUSIONS

## Conclusions

A high-density SNP genetic map was constructed for cultivated tetraploid cotton, using an inter-specific RIL population. Over 1 million SNPs were used to create the genetic map, covering all 26 cotton chromosomes. Using the genetic map and several years of phenotypic data for lint percent and seed index, 18 potential QTLs were able to be mapped with explained phenotypic variation ranging from 6.15 to 37.48%. This map is one of the first maps created using GBS/ddRADseq technology. It is also one of the longest and most dense genetic maps to date. Large GxE effects and gene-gene interactions make it difficult, if not impossible, to identify the same QTL in every environment. Smaller geographical areas or multiple locations with similar environments might be best suited for breeding with MAS. The results of this study not only further the research behind the implementation of MAS into cotton breeding programs, but also help realize the limitations of MAS in its application to large geographical areas/ different environments.

## REFERENCES


Abdalla, A., O. Reddy, K. El-Zik and A. Pepper. 2001. Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. Theor. Appl. Genet. 102: 222-229.

An, C., J.N. Jenkins, J. Wu, Y. Guo and J.C. McCarty. 2009. Use of fiber and fuzz mutants to detect QTL for yield components, seed, and fiber traits of upland cotton. Euphytica 172: 21-34. doi:10.1007/s10681-009-0009-2.

An, C., S. Saha, J.N. Jenkins, D.P. Ma, B.E. Scheffler, R.J. Kohel, et al. 2008. Cotton (Gossypium spp.) R2R3-MYB transcription factors SNP identification, phylogenomic characterization, chromosome localization, and linkage mapping. Theor. Appl. Genet. 116: 1015-1026. doi:10.1007/s00122-008-0732-4.

An, C., S. Saha, J.N. Jenkins, B.E. Scheffler, T.A. Wilkins and D.M. Stelly. 2007. Transcriptome profiling, sequence characterization, and SNP-based chromosomal assignment of the EXPANSIN genes in cotton. Mol. Genet. Genomics 278: 539-553. doi:10.1007/s00438-007-0270-9.

Bartlett, J.M.S. and D. Stirling. 2003. A Short History of the Polymerase Chain Reaction. In: J. M. S. Bartlett and D. Stirling, editors, PCR Protocols. Humana Press, Totowa, NJ. p. 3-6.1994. The power and deceit of QTL experiments: lessons from comparative QTL studies. Proceedings of the forty-ninth annual corn and sorghum industry research conference, Chicago, IL.

Batley, J., G. Barker, H. O'Sullivan, K.J. Edwards, and D. Edwards. 2003. Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. Plan Physiol., *132*(1), 84-91.

Beavis, W.D. 1998. QTL analyses: power, precision, and accuracy. Molecular Dissection of Complex Traits 1998: 145-162.

Blenda, A., J. Scheffler, B. Scheffler, M. Palmer, J.M. Lacape, J.Z. Yu, et al. 2006. CMD: a Cotton Microsatellite Database resource for Gossypium genomics. BMC Genomics 7: 132. doi:10.1186/1471-2164-7-132.

Brubaker, C.L., A. Paterson and J.F. Wendel. 1999. Comparative genetic mapping of allotetraploid cotton and its diploid progenitors. Genome 42: 184-203.

Brubaker, C.L., F. Bourland and J.F. Wendel. 1999. The origin and domestication of cotton. Cotton: Origin, History, Technology, and Production. John Wiley & Sons, New York: 3-31.

Brubaker, C.L. and A.H. Brown. 2003. The use of multiple alien chromosome addition aneuploids facilitates genetic linkage mapping of the Gossypium G genome. Genome 46: 774-791. doi:10.1139/g03-063.

Brubaker, C.L. and J.F. Wendel. 1994. Reevaluating the origin of domesticated cotton (Gossypium hirsutum; Malvaceae) using nuclear restriction fragment length polymorphisms (RFLPs). Am. J. Bot. 1309-1326.

Bundock, P.C. and R.J. Henry. 2004. Single nucleotide polymorphism, haplotype diversity and recombination in the Isa gene of barley. Theor. Appl. Genet. 109: 543-551.

Burr, B., F.A. Burr, K.H. Thompson, M. Albertson and C.W. Stuber. 1988. Gene mapping with recombinant inbreds in maize. Genetics 118: 519-526.

Byers, R.L., D.B. Harker, S.M. Yourstone, P.J. Maughan and J.A. Udall. 2012. Development and mapping of SNP assays in allotetraploid cotton. Theor. Appl. Genet. 124: 1201-1214. doi:10.1007/s00122-011-1780-8.

Caldwell, K.S., J. Dvorak, E.S. Lagudah, E. Akhunov, M.-C. Luo, P. Wolters, et al. 2004. Sequence polymorphism in polyploid wheat and their D-genome diploid ancestor. Genetics 167: 941-947.

Chaparro, J., D. Werner, D. O'malley and R. Sederoff. 1994. Targeted mapping and linkage analysis of morphological isozyme, and RAPD markers in peach. Theor. Appl. Genet. 87: 805-815.

Chee, P., X. Draye, C.-X. Jiang, L. Decanini, T.A. Delmonte, R. Bredhauer, et al. 2005. Molecular dissection of interspecific variation between Gossypium hirsutum and Gossypium barbadense (cotton) by a backcross-self approach: I. Fiber elongation. Theor. Appl. Genet. 111: 757-763.

Chee, P.W., X. Draye, C.-X. Jiang, L. Decanini, T.A. Delmonte, R. Bredhauer, et al. 2005. Molecular dissection of phenotypic variation between Gossypium hirsutum and Gossypium barbadense (cotton) by a backcross-self approach: III. Fiber length. Theor. Appl. Genet. 111: 772-781.

Chen, H., N. Qian, W. Guo, Q. Song, B. Li, F. Deng, et al. 2009. Using three overlapped RILs to dissect genetically clustered QTL for fiber strength on Chro. D8 in Upland cotton. Theor. Appl. Genet. 119: 605-612.

Ching, A., K.S. Caldwell, M. Jung, M. Dolan, O. Smith, S. Tingey, et al. 2002. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. BMC Genet. 3: 1.

Collard, B.C. and D.J. Mackill. 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 363: 557-572. doi:10.1098/rstb.2007.2170.

Draye, X., P. Chee, C.-X. Jiang, L. Decanini, T.A. Delmonte, R. Bredhauer, et al. 2005. Molecular dissection of interspecific variation between Gossypium hirsutum and G. barbadense (cotton) by a backcross-self approach: II. Fiber fineness. Theor. Appl. Genet. 111: 764-771.

Edwards, D., J.W. Forster, D. Chagné and J. Batley. 2007. What Are SNPs? Association Mapping in Plants. Springer. p. 41-52.

Eldredge, L., R. Ballard, W. Baird, A. Abbott, P. Morgens, A. Callahan, et al. 1992. Application of RFLP analysis to genetic linkage mapping in peaches. Hort. Science 27: 160-163.

Erickson, D.L., C.B. Fenster, H.K. Stenoien and D. Price. 2004. Quantitative trait locus analyses and the study of evolutionary process. Mol. Ecol. 13: 2505-2522. doi:10.1111/j.1365-294X.2004.02254.x.

Fang, D.D., J. Xiao, P.C. Canci and R.G. Cantrell. 2010. A new SNP haplotype associated with blue disease resistance gene in cotton (Gossypium hirsutum L.). Theor. Appl. Genet. 120: 943-953. doi:10.1007/s00122-009-1223-y.

Feltus, F.A., J. Wan, S.R. Schulze, J.C. Estill, N. Jiang and A.H. Paterson. 2004. An SNP resource for rice genetics and breeding based on subspecies indica and japonica genome alignments. Genome Res. 14: 1812-1819.

Ferreira, A., M.F.d. Silva and C.D. Cruz. 2006. Estimating the effects of population size and type on the accuracy of genetic maps. Genet. Mol. Biol. 29: 187-192.

Frelichowski Jr, J.E., M.B. Palmer, D. Main, J.P. Tomkins, R.G. Cantrell, D.M. Stelly, et al. 2006. Cotton genome mapping with new microsatellites from Acala 'Maxxa'BAC-ends. Mol. Genet. Genomics 275: 479-491.

Gore, M.A., D.D. Fang, J.A. Poland, J. Zhang, R.G. Percy, R.G. Cantrell, et al. 2014. Linkage Map Construction and Quantitative Trait Locus Analysis of Agronomic and Fiber Quality Traits in Cotton. Plant Genome 7: 0. doi:10.3835/plantgenome2013.07.0023.

Grandillo, S. and S. Tanksley. 1996. QTL analysis of horticultural traits differentiating the cultivated tomato from the closely related species Lycopersicon pimpinellifolium. Theor. Appl. Genet. 92: 935-951.

Guo, W., C. Cai, C. Wang, Z. Han, X. Song, K. Wang, et al. 2007. A microsatellite-based, gene-rich linkage map reveals genome structure, function and evolution in Gossypium. Genetics 176: 527-541. doi:10.1534/genetics.107.070375.

Huang, N., S. McCouch, T. Mew, A. Parco and E. Guiderdoni. 1994. 33. Development of an RFLP Map from a Doubled Haploid Population in Rice.

Hulse-Kemp, A.M., H. Ashrafi, X. Zheng, F. Wang, K.A. Hoegenauer, A.B. Maeda, et al. 2014. Development and bin mapping of gene-associated interspecific SNPs for cotton (Gossypium hirsutum L.) introgression breeding efforts. BMC Genomics 15: 945. doi:10.1186/1471-2164-15-945.

Iqbal, M.J., N. Aziz, N. Saeed, Y. Zafar and K. Malik. 1997. Genetic diversity evaluation of some elite cotton varieties by RAPD analysis. Theor. Appl. Genet. 94: 139-144.

Jarne, P. and P.J. Lagoda. 1996. Microsatellites, from molecules to populations and back. Trends Ecol. Evol. 11: 424-429.

Jiang, C.-x., R. Wright, S. Woo, T. DelMonte and A. Paterson. 2000. QTL analysis of leaf morphology in tetraploid Gossypium (cotton). Theor. Appl. Genet. 100: 409-418.

Jiang, C.-X., R.J. Wright, K.M. El-Zik and A.H. Paterson. 1998. Polyploid formation created unique avenues for response to selection in Gossypium (cotton). Proc. Natl. Acad. Sci. 95: 4419-4424.

Johnson, J., MacDonald, S., Meyer, L., Norrington B., Skelly, C. 2014. The World and United States Cotton Outlook. Agricultural Outlook Forum 2014. U.S. Department of Agriculture.

Kanazin, V., H. Talbert, D. See, P. DeCamp, E. Nevo and T. Blake. 2002. Discovery and assay of single-nucleotide polymorphisms in barley (Hordeum vulgare). Plant Mol. Biol. 48: 529-537.

Kesseli, R.V., I. Paran and R.W. Michelmore. 1994. Analysis of a detailed genetic linkage map of Lactuca sativa (lettuce) constructed from RFLP and RAPD markers. Genetics 136: 1435-1446.

Khan, S.A., D. Hussain, E. Askari, J.M. Stewart, K. Malik and Y. Zafar. 2000. Molecular phylogeny of Gossypium species by DNA fingerprinting. Theor. Appl. Genet. 101: 931-938.

Kim, M., K. Van, P. Lestari, J.-K. Moon and S.-H. Lee. 2005. SNP identification and SNAP marker development for a GmNARK gene controlling supernodulation in soybean. Theor. Appl. Genet. 110: 1003-1010.

Kohel, R.J., J. Yu, Y.-H. Park and G.R. Lazo. 2001. Molecular mapping and characterization of traits controlling fiber quality in cotton. Euphytica 121: 163-172.

Lacape, J.-M., T.-B. Nguyen, B. Courtois, J.-L. Belot, M. Giband, J.-P. Gourlot, et al. 2005. QTL analysis of cotton fiber quality using multiple× backcross generations. Crop Sci. 45: 123-140.

Lacape, J.-M., T.-B. Nguyen, S. Thibivilliers, B. Bojinov, B. Courtois, R. Cantrell, et al. 2003. A combined RFLP SSR AFLP map of tetraploid cotton based on a Gossypium hirsutum× Gossypium barbadense backcross population. Genome 46: 612-626.

Lacape J. M., Jacobs J., Arioli T., Derijcker R., Forestier-Chiron N., et al. , 2009. A new interspecific, *Gossypium hirsutum* x *G. barbadense*, RIL population: towards a unified consensus linkage map of tetraploid cotton. Theor. Appl. Genet. 119: 281–292

Lee, M.-K., Y. Zhang, M. Zhang, M. Goebel, H.J. Kim, B.A. Triplett, et al. 2013. Construction of a plant-transformation-competent BIBAC library and genome sequence analysis of polyploid Upland cotton (Gossypium hirsutum L.). BMC genomics 14: 1.

Li, F., G. Fan, C. Lu, G. Xiao, C. Zou, R.J. Kohel, et al. 2015. Genome sequence of cultivated Upland cotton (Gossypium hirsutum TM-1) provides insights into genome evolution. Nat. Biotechnol. 33: 524-530. doi:10.1038/nbt.3208.

Lin, Z., Y. Zhang, X. Zhang and X. Guo. 2008. A high-density integrative linkage map for Gossypium hirsutum. Euphytica 166: 35-45. doi:10.1007/s10681-008-9822-2.

Lin, Z.-x., D. He, X.-l. Zhang, Y. Nie, X. Guo, C. Feng, et al. 2005. Linkage map construction and mapping QTL for cotton fibre quality using SRAP, SSR and RAPD. Plant Breed. 124: 180-187.

Liu, Y-H. 2014. Molecular basis of quantitative genetics revealed by cloning and analysis of 474 genes controlling fiber length in cotton. Ph.D. dissertation, Texas A&M University, College Station, Texas.

Liu, Z.-W., Biyashev, R. M., and Maroof, M. A. S. 1996 Development of simple sequence repeat DNA markers and their integration into a barley linkage map. Theor. Appl. Genet., *93*(5), 869–876. http://doi.org/10.1007/BF00224088

Liu, S., S. Saha, D. Stelly, B. Burr, and R.G. Cantrell. 2000. Chromosomal assignment of microsatellite loci in cotton. J. Hered. 91: 326-332.

Lu, Y., J. Curtiss, R.G. Percy, S.E. Hughs, S. Yu and J. Zhang. 2009. DNA Polymorphisms of Genes Involved in Fiber Development in a Selected Set of Cultivated Tetraploid Cotton. Crop Sci. 49: 1695. doi:10.2135/cropsci2008.12.0693.

Maheswaran, M., P. Subudhi, S. Nandi, J. Xu, A. Parco, D. Yang, et al. 1997. Polymorphism, distribution, and segregation of AFLP markers in a doubled haploid rice population. Theor. Appl. Genet. 94: 39-45.

Maughan, P., M.S. Maroof, G. Buss and G. Huestis. 1996. Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near-isogenic line analysis. Theor. Appl. Genet. 93: 392-401.

Mei, M., N.H. Syed, W. Gao, P.M. Thaxton, C.W. Smith, D.M. Stelly, et al. 2004. Genetic mapping and QTL analysis of fiber-related traits in cotton ( Gossypium). Theor. Appl. Genet. 108: 280-291. doi:10.1007/s00122-003-1433-7.

Meksem, K., D. Leister, J. Peleman, M. Zabeau, F. Salamini and C. Gebhardt. 1995. A high-resolution map of the vicinity of the R1 locus on chromosome V of potato based on RFLP and AFLP markers. Mol. Gen. Genet. MGG 249: 74-81.

Meng, L., H. Li, L. Zhang, and J. Wang. 2015. IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. Crop J. 3(3): 269-283.

Meredith W.R. 1992. RFLP association with varietal origin and heterosis. In : Herber D (ed) Proc Beltwide Cotton Prod Res Conf. Nashville, TN, 6–10 Jan. National Cotton Council of America, Memphis, Tennessee, p 607

Meudt, H.M. and A.C. Clarke. 2007. Almost forgotten or latest practice? AFLP applications, analyses and advances. Trends Plant Sci. 12: 106-117.

Mochida, K., Y. Yamazaki and Y. Ogihara. 2004. Discrimination of homoeologous gene expression in hexaploid wheat by SNP analysis of contigs grouped from a large number of expressed sequence tags. Mol. Genet. Genomics 270: 371-377.

Mueller, U.G. and L.L. Wolfenbarger. 1999. AFLP genotyping and fingerprinting. Trends Ecol. Evol. 14: 389-394.

Multani, D. and B. Lyon. 1995. Genetic fingerprinting of Australian cotton cultivars with RAPD markers. Genome 38: 1005-1008.

Murray, M., J. Cramer, Y. Ma, D. West, J. Romero-Severson, J. Pitas, et al. 1988. Agrigenetics maize RFLP linkage map. Maize Genetics Cooperation Newsletter 62: 89.

Paterson, A.H., C.L. Brubaker and J.F. Wendel. 1993. A rapid method for extraction of cotton (Gossypium spp.) genomic DNA suitable for RFLP or PCR analysis. Plant Mol. Biol. Rep. 11: 122-127.

Paterson, A.H., Y. Saranga, M. Menz, C.X. Jiang and R.J. Wright. 2003. QTL analysis of genotype x environment interactions affecting cotton fiber quality. Theor. Appl. Genet. 106: 384-396. doi:10.1007/s00122-002-1025-y.

Peterson, B.K., J.N. Weber, E.H. Kay, H.S. Fisher and H.E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One 7: e37135. doi:10.1371/journal.pone.0037135.

Qin, H., M. Chen, X. Yi, S. Bie, C. Zhang, Y. Zhang, et al. 2015. Identification of associated SSR markers for yield component and fiber quality traits based on frame map and Upland cotton collections. PLoS One 10: e0118073. doi:10.1371/journal.pone.0118073.

Rami, J.-F., P. Dufour, G. Trouche, G. Fliedel, C. Mestres, F. Davrieux, et al. 1998. Quantitative trait loci for grain quality, productivity, morphological and agronomical traits in sorghum (Sorghum bicolor L. Moench). Theor. Appl. Genet. 97: 605-616.

Rana, M. and K. Bhat. 2004. A comparison of AFLP and RAPD markers for genetic diversity and cultivar identification in cotton. J. Plant Biochem. Biotechnol. 13: 19-24.

Reinisch, A.J., J.-M. Dong, C.L. Brubaker, D.M. Stelly, J.F. Wendel and A.H. Paterson. 1994. A detailed RFLP map of cotton, Gossypium hirsutum x Gossypium barbadense: chromosome organization and evolution in a disomic polyploid genome. Genetics 138: 829-847.

Reiter, R.S., J. Williams, K.A. Feldmann, J.A. Rafalski, S.V. Tingey and P.A. Scolnik. 1992. Global and local genome mapping in Arabidopsis thaliana by using recombinant inbred lines and random amplified polymorphic DNAs. Proc. Natl. Acad. Sci. 89: 1477-1481.

Roberts, C., R.G. Cantrell, and S.T. Ball. 1997. Release of New Mexico Sea Island 1331. p. 473 In Proc. Beltwide Cotton Conf., New Orleans, LA. 7-10 Jan. 1997. Natl. Cotton Counc. Am., Memphis, TN.

Rong, J., C. Abbey, J.E. Bowers, C.L. Brubaker, C. Chang, P.W. Chee, et al. 2004. A 3347-locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission and evolution of cotton (Gossypium). Genetics 166: 389-417.

Rong, J., F.A. Feltus, V.N. Waghmare, G.J. Pierce, P.W. Chee, X. Draye, et al. 2007. Meta-analysis of polyploid cotton QTL shows unequal contributions of subgenomes to a complex network of genes and gene clusters implicated in lint fiber development. Genetics 176: 2577-2588. doi:10.1534/genetics.107.074518.

SAS Institute Inc. 2013. Base SAS® 9.4 Procedures Guide OnlineDoc. Cary, NC: SAS Institute Inc.

Shappley, Z.W. 1996. Construction of RFLP linkage groups and mapping of quantitative trait loci in upland cotton(Gossypium hirsutum L.). Ph.D. diss. Mississippi State Univ., Mississippi State, (Diss. Abstr. AAG9711761).

Shappley, Z., J.N. Jenkins, J. Zhu and J.C. McCarty Jr. 1998. Quantitative trait loci associated with agronomic and fiber traits of upland cotton. J. Cotton Sci.

Shappley, Z.W., J.N. Jenkins, W.R. Meredith and J.C. McCarty Jr. 1998. An RFLP linkage map of Upland cotton, Gossypium hirsutum L. Theor. Appl. Genet. 97: 756-761.

Shen, X., W. Guo, X. Zhu, Y. Yuan, J.Z. Yu, R.J. Kohel, et al. 2005. Molecular mapping of QTLs for fiber qualities in three diverse lines in Upland cotton using SSR markers. Mol. Breed. 15: 169-181. doi:10.1007/s11032-004-4731-0.

Shen, X., T. Zhang, W. Guo, X. Zhu and X. Zhang. 2006. Mapping Fiber and Yield QTLs with Main, Epistatic, and QTL × Environment Interaction Effects in Recombinant Inbred Lines of Upland Cotton. Crop Sci. 46: 61. doi:10.2135/cropsci2005.0056.

Small, R. and J. Wendel. 1998. The mitochondrial genome of allotetraploid cotton (Gossypium L.). J. Hered. 90: 251-253.

Smith C. 2003. Registration of TAM 94L-25 and TAM 94J-3 germplasm lines of upland cotton with improved fiber length. Crop Sci. 43:742-743.

Somers, D.J., R. Kirkpatrick, M. Moniwa and A. Walsh. 2003. Mining single-nucleotide polymorphisms from hexaploid wheat ESTs. Genome 46: 431-437.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503-517.

Tang, S., Z. Teng, T. Zhai, X. Fang, F. Liu, D. Liu, et al. 2014. Construction of genetic map and QTL analysis of fiber quality traits for Upland cotton (Gossypium hirsutum L.). Euphytica 201: 195-213. doi:10.1007/s10681-014-1189-y.

Tatineni, V., R. Cantrell and D. Davis. 1996. Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. Crop Sci. 36: 186-192.

Tomkins, J., D. Peterson, T. Yang, D. Main, T. Wilkins, A. Paterson, et al. 2001. Development of genomic resources for cotton (Gossypium hirsutum L.): BAC library construction, preliminary STC analysis, and identification of clones associated with fiber development. Mol. Breed. 8: 255-261.

Ulloa, M., R. Cantrell, R. Percy, E. Zeiger and L. ZhenMin. 2000. QTL analysis of stomatal conductance and relationship to lint yield in an interspecific cotton. J. Cotton Sci. 4: 10-18.

Ulloa, M. and W.R. Meredith Jr. 2000. Genetic linkage map and QTL analysis of agronomic and fiber quality traits in an intraspecific population. J. Cotton Sci. 4: 161-170.

Ulloa, M., S. Saha, J. Jenkins, W. Meredith, J. McCarty and D. Stelly. 2005. Chromosomal assignment of RFLP linkage groups harboring important QTLs on an intraspecific cotton (Gossypium hirsutum L.) joinmap. J. Hered. 96: 132-144.

USDA. 2016. Oilseeds: World Markets and Trade. United States Department of Agriculture. Retrieved January 25, 2016. (http://usda.mannlib.cornell.edu/usda/current/oilseed-trade/oilseed-trade-01-12-2016.pdf).

US Climate Data. 2016. Version 2.2. Retrieved March 7, 2016. (http://www.usclimatedata.com/)

Van Deynze, A., K. Stoffel, M. Lee, T.A. Wilkins, A. Kozik, R.G. Cantrell, et al. 2009. Sampling nucleotide diversity in cotton. BMC Plant Biol. 9: 125. doi:10.1186/1471-2229-9-125.

Varshney, R.K., S.N. Nayak, G.D. May and S.A. Jackson. 2009. Next-generation sequencing technologies and their implications for crop genetics and breeding. Trends Biotechnol. 27: 522-530. doi:10.1016/j.tibtech.2009.05.006.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Van de Lee, M. Hornes, et al. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23: 4407-4414.

Wang, H., C. Huang, H. Guo, X. Li, W. Zhao, B. Dai, et al. 2015. QTL Mapping for Fiber and Yield Traits in Upland Cotton under Multiple Environments. PLoS One 10: e0130742. doi:10.1371/journal.pone.0130742.

Wang, K., Z. Wang, F. Li, W. Ye, J. Wang, G. Song, et al. 2012. The draft genome of a diploid cotton Gossypium raimondii. Nat. Genet. 44: 1098-1103. doi:10.1038/ng.2371.

Wang, S., J. Chen, W. Zhang, Y. Hu, L. Chang, L. Fang, et al. 2015. Sequence-based ultra-dense genetic and physical maps reveal structural variations of allopolyploid cotton genomes. Genome Biol. 16: 108. doi:10.1186/s13059-015-0678-1.

Wang, Y., Z. Ning, Y. Hu, J. Chen, R. Zhao, H. Chen, et al. 2015. Molecular Mapping of Restriction-Site Associated DNA Markers In Allotetraploid Upland Cotton. PLoS One 10: e0124781. doi:10.1371/journal.pone.0124781.

Wang, Z., G. Second and S. Tanksley. 1992. Polymorphism and phylogenetic relationships among species in the genus Oryza as determined by analysis of nuclear RFLPs. Theor. Appl. Genet. 83: 565-581.

Wendel, J.F. and V.A. Albert. 1992. Phylogenetics of the cotton genus (Gossypium): character-state weighted parsimony analysis of chloroplast-DNA restriction site data and its systematic and biogeographic implications. Syst. Bot. 115-143.

Wendel, J.F., P.D. Olson and J. McD. 1989. Genetic diversity, introgression, and independent domestication of old world cultivated cottons. Am. J. Bot. 1795-1806.

Williams, J.G., A.R. Kubelik, K.J. Livak, J.A. Rafalski and S.V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18: 6531-6535.

Winter, P. and G. Kahl. 1995. Molecular marker technologies for plant improvement. World J. Microbiol. Biotechnol. 11: 438-448.

Xiao, J., D.D. Fang, M. Bhatti, B. Hendrix and R. Cantrell. 2009. A SNP haplotype associated with a gene resistant to Xanthomonas axonopodis pv. malvacearum in upland cotton (Gossypium hirsutum L.). Mol. Breed. 25: 593-602. doi:10.1007/s11032-009-9355-y.

Xiao, J., K. Wu, D.D. Fang, D.M. Stelly, J. Yu and R.G. Cantrell. 2009. New SSR markers for use in cotton (Gossypium spp.) improvement. J. Cotton Sci. 13: 75-157.

Yang, S., P. Latreille, D. Zijin, D. Caldwell, R. Kerstetter. 2015. TM-1 Sequence Data. Monsanto Company, 800 N. Lindbergh Blvd., St. Louis, Missouri. Personal communication.

Yu, J., S. Yu, C. Lu, W. Wang, S. Fan, M. Song, et al. 2007. High-density linkage map of cultivated allotetraploid cotton based on SSR, TRAP, SRAP and AFLP markers. J. of Integr. Plant Biol. 49: 716-724.

Yu, J., K. Zhang, S. Li, S. Yu, H. Zhai, M. Wu, et al. 2013. Mapping quantitative trait loci for lint yield and fiber quality across environments in a Gossypium hirsutum x

Gossypium barbadense backcross inbred line population. Theor. Appl. Genet. 126: 275-287. doi:10.1007/s00122-012-1980-x.

Yu, J.Z., R.J. Kohel, D.D. Fang, J. Cho, A. Van Deynze, M. Ulloa, et al. 2012. A high-density simple sequence repeat and single nucleotide polymorphism genetic map of the tetraploid cotton genome. G3 (Bethesda) 2: 43-58. doi:10.1534/g3.111.001552.

Yu, S., K. Wang, F. Li, R.J. Kohel, R.G. Percy and J. Yu. 2008. Sequencing of the cultivated tetraploid cotton genome-Gossypium hirsutum. Cotton Sci. 20: 4.

Yu, Y., D. Yuan, S. Liang, X. Li, X. Wang, Z. Lin, et al. 2011. Genome structure of cotton revealed by a genome-wide SSR genetic map constructed from a BC 1 population between Gossypium hirsutum and G. barbadense. BMC Genomics 12: 1.

Yu, J. 2009. The Construction of a Gossypium AD-genome-wide Comprehensive Reference Map Based on Diverse Data Resources (Doctoral dissertation, Texas A&M University).

Zhang, J., W. Guo and T. Zhang. 2002. Molecular linkage map of allotetraploid cotton (Gossypium hirsutum L.× Gossypium barbadense L.) with a haploid population. Theor. Appl. Genet. 105: 1166-1174.

Zhang, J., Y. Lu and S. Yu. 2005. Cleaved AFLP (cAFLP), a modified amplified fragment length polymorphism analysis for cotton. Theor. Appl. Genet. 111: 1385-1395.

Zhang, K., J. Zhang, J. Ma, S. Tang, D. Liu, Z. Teng, et al. 2012. Genetic mapping and quantitative trait locus analysis of fiber quality traits using a three-parent composite population in upland cotton (Gossypium hirsutum L.). Mol. Breed. 29: 335-348.

Zhang, T., Y. Hu, W. Jiang, L. Fang, X. Guan, J. Chen, et al. 2015. Sequencing of allotetraploid cotton (Gossypium hirsutum L. acc. TM-1) provides a resource for fiber improvement. Nat. Biotechnol. 33: 531-537. doi:10.1038/nbt.3207.

Zhang, W., M. Gianibelli, W. Ma, L. Rampling and K. Gale. 2003. Identification of SNPs and development of allele-specific PCR markers for γ-gliadin alleles in Triticum aestivum. Theor. Appl. Genet. 107: 130-138.

Zhang, Z.-S., M.-C. Hu, J. Zhang, D.-J. Liu, J. Zheng, K. Zhang, et al. 2009. Construction of a comprehensive PCR-based marker linkage map and QTL mapping for fiber quality traits in upland cotton (Gossypium hirsutum L.). Mol. Breed. 24: 49-61. doi:10.1007/s11032-009-9271-1.

Zhao, L., L. Yuanda, C. Caiping, T. Xiangchao, C. Xiangdong, Z. Wei, et al. 2012. Toward allotetraploid cotton genome assembly: integration of a high-density molecular genetic linkage map with DNA sequence information. BMC Genomics 13: 539. doi:10.1186/1471-2164-13-539.

Zhu, Y., Q. Song, D. Hyten, C. Van Tassell, L. Matukumalli, D. Grimm, et al. 2003. Single-nucleotide polymorphisms in soybean. Genetics 163: 1123-1134.

## APPENDIX

**Table A-1**. Raw phenotypic data for lint percent and seed index of the RIL population in all environments.

| | Lint Percent | | | | | | Seed Index | | |
|---|---|---|---|---|---|---|---|---|---|
| Entry | 2009 CS | 2010 CS | 2011 CS | 2015 LA | 2015 CS | 2015 LU | 2015 LA | 2015 CS | 2015 LU |
| 1 | 39.00 | - | 32.67 | 33.33 | 34.18 | 36.08 | 13.11 | 12.23 | 11.92 |
| 2 | 35.00 | 40.00 | 29.33 | 31.99 | 34.47 | 38.02 | 13.08 | 11.37 | 10.72 |
| 3 | 37.00 | 35.67 | 29.67 | 37.89 | 37.09 | 41.68 | 11.25 | 10.30 | 10.72 |
| 4 | 45.00 | 40.00 | 31.00 | 33.64 | 35.51 | 39.99 | 13.96 | 11.38 | 11.52 |
| 5 | 32.00 | - | 31.00 | 34.78 | 34.45 | 38.63 | 10.51 | 9.86 | 10.14 |
| 6 | 35.00 | 33.67 | 31.00 | 35.35 | 36.70 | 40.79 | 12.65 | 10.54 | 10.80 |
| 7 | 31.00 | 41.50 | 30.33 | 34.33 | 34.78 | 39.90 | 11.01 | 11.27 | 9.71 |
| 8 | 37.00 | - | 33.33 | 34.21 | 34.72 | 36.44 | 11.55 | 10.12 | 10.57 |
| 9 | 30.00 | 26.50 | 25.00 | 34.32 | 33.73 | 38.08 | 12.70 | 11.00 | 11.02 |
| 10 | 24.00 | 25.00 | 24.33 | 36.94 | 36.33 | 38.21 | 11.81 | 11.77 | 11.57 |
| 11 | 34.00 | - | 31.33 | 35.87 | 37.88 | 39.86 | 12.06 | 10.34 | 10.76 |
| 12 | 42.00 | 45.33 | 41.33 | 34.44 | 35.71 | 37.26 | 12.64 | 11.21 | 11.63 |
| 13 | 18.00 | 18.00 | 18.33 | 37.06 | 36.25 | 42.58 | 12.31 | 12.03 | 11.31 |
| 14 | 35.00 | 34.00 | 35.00 | 34.41 | 35.34 | 39.68 | 13.22 | 10.86 | 11.75 |
| 15 | 34.00 | - | 29.00 | 36.26 | 36.81 | 39.60 | 13.77 | 11.97 | 13.13 |
| 16 | 35.00 | 33.67 | 27.67 | 32.68 | 35.68 | 37.45 | 11.89 | 10.36 | 10.76 |
| 17 | 45.00 | 38.00 | 35.50 | - | 36.38 | 40.06 | - | 10.64 | 10.78 |
| 18 | 38.00 | 34.50 | 32.50 | 34.31 | 36.72 | 41.75 | 11.85 | 10.10 | 10.48 |
| 20 | 33.00 | - | 28.00 | 34.35 | 35.33 | 38.51 | 12.96 | 11.26 | 11.24 |
| 21 | 32.00 | 34.67 | 32.33 | 33.08 | 33.67 | 35.60 | 12.81 | 10.52 | 11.13 |
| 22 | 39.00 | 44.50 | 36.33 | 37.17 | 36.33 | 40.00 | 10.23 | 10.02 | 10.75 |
| 23 | 33.00 | 37.67 | 36.67 | 36.61 | 36.74 | 38.98 | 11.76 | 10.56 | 11.52 |
| 24 | 29.00 | 28.00 | 25.67 | 33.50 | 33.76 | 38.91 | 12.49 | 11.29 | 11.89 |
| 25 | 38.00 | 35.00 | 32.33 | 33.99 | 36.35 | 39.55 | 12.65 | 11.00 | 10.97 |
| 30 | 32.00 | 36.67 | 32.67 | 32.89 | 34.45 | 40.59 | 12.91 | 10.64 | 12.20 |
| 31 | 34.00 | - | 30.00 | 32.84 | 33.77 | 38.32 | 15.78 | 12.47 | 11.95 |
| 32 | 24.00 | 26.67 | 26.33 | 27.17 | 29.92 | 35.58 | 14.79 | 12.89 | 12.05 |
| 33 | 35.00 | 43.00 | 34.33 | 35.49 | 34.67 | 38.35 | 10.34 | 9.40 | 11.52 |
| 34 | 29.00 | . | 26.00 | 33.96 | 36.32 | 37.94 | 13.39 | 11.36 | 11.95 |
| 35 | - | 33.00 | 31.33 | 34.34 | 35.11 | 40.76 | 12.03 | 11.21 | 11.43 |
| 38 | 32.00 | 39.33 | 31.33 | 33.09 | 35.38 | 40.74 | 12.59 | 10.51 | 10.06 |
| 39 | 36.00 | - | 31.67 | 35.64 | 34.44 | 38.42 | 12.03 | 10.26 | 10.76 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 40 | 37.00 | 37.00 | 33.33 | 38.74 | 37.83 | 40.63 | 11.37 | 10.30 | 10.66 |
| 41 | 32.00 | 31.67 | 31.33 | 30.80 | 32.97 | 36.94 | 12.47 | 10.73 | 10.79 |
| 42 | 31.00 | 32.33 | 31.00 | 33.01 | 35.89 | 39.73 | 11.44 | 9.98 | 11.46 |
| 43 | 26.00 | 32.00 | 30.00 | 31.89 | 35.18 | 39.36 | 12.78 | 10.63 | 11.14 |
| 44 | 45.00 | 46.00 | 41.00 | 38.72 | 42.56 | 41.72 | 11.91 | 9.18 | 10.67 |
| 45 | 42.00 | 44.33 | 38.67 | 39.68 | 40.74 | 43.81 | 10.53 | 8.74 | 10.90 |
| 46 | 27.00 | - | 26.67 | 32.70 | 33.43 | 36.39 | 13.43 | 11.35 | 12.81 |
| 47 | 34.00 | 38.00 | 35.33 | 33.36 | 35.55 | 40.52 | 11.77 | 9.97 | 11.28 |
| 48 | 31.00 | - | 24.67 | 33.44 | 35.21 | 37.97 | 12.93 | 9.84 | 9.86 |
| 49 | 33.00 | 30.00 | 28.67 | 34.51 | 33.37 | 41.82 | 11.32 | 10.02 | 11.67 |
| 50 | 35.00 | 30.00 | 29.33 | 27.85 | 30.13 | 36.60 | 15.63 | 11.87 | 11.88 |
| 51 | 35.00 | 33.00 | 34.33 | 33.42 | 35.53 | 37.02 | 12.91 | 10.22 | 10.19 |
| 52 | 40.00 | - | 36.33 | 35.01 | 36.47 | 39.39 | 12.16 | 10.47 | 10.55 |
| 53 | 38.00 | 37.00 | 35.33 | 37.10 | 38.11 | 40.59 | 13.30 | 10.31 | 10.30 |
| 54 | 25.00 | 26.00 | 25.00 | 30.64 | 31.19 | 34.95 | 14.59 | 11.50 | 12.54 |
| 59 | 43.00 | 33.33 | 38.67 | 38.78 | 39.21 | 41.88 | 11.35 | 9.86 | 10.76 |
| 60 | 37.00 | 37.50 | 32.67 | 32.36 | 33.57 | 36.96 | 13.26 | 11.24 | 11.25 |
| 61 | 41.00 | - | 35.50 | 36.48 | 37.02 | 43.60 | 11.78 | 9.90 | 10.38 |
| 62 | 26.00 | 20.00 | 18.33 | 32.52 | 31.14 | 37.80 | 12.34 | 11.31 | 10.76 |
| 63 | 45.00 | 32.50 | 31.67 | 33.54 | 32.13 | 38.29 | 13.64 | 11.05 | 11.34 |
| 64 | 33.00 | 29.00 | 26.33 | 31.22 | 33.02 | 37.75 | 13.73 | 11.02 | 10.49 |
| 65 | 40.00 | 39.33 | 37.00 | 36.20 | 34.59 | 33.84 | 13.47 | 11.41 | 8.61 |
| 66 | 34.00 | 30.33 | 30.00 | 32.19 | 32.78 | 38.11 | 13.98 | 11.35 | 12.61 |
| 67 | 35.00 | 32.00 | 32.67 | 32.83 | 34.97 | 35.25 | 13.63 | 10.93 | 11.17 |
| 68 | 36.00 | 33.50 | 29.67 | 34.47 | 32.98 | 40.39 | 13.46 | 11.55 | 10.82 |
| 69 | 30.00 | 24.67 | 26.67 | 30.71 | 29.97 | 34.17 | 12.94 | 11.59 | 10.92 |
| 70 | 28.00 | 24.00 | 26.67 | 31.34 | 33.58 | 36.78 | 14.68 | 12.49 | 13.69 |
| 71 | 47.00 | 28.50 | 26.67 | 32.38 | 33.50 | 38.54 | 12.95 | 11.55 | 11.57 |
| 72 | 35.00 | - | 29.67 | 35.35 | 36.54 | 40.71 | 13.16 | 10.74 | 11.18 |
| 73 | 39.00 | 40.33 | 36.67 | 33.77 | 34.83 | 38.25 | 15.36 | 11.85 | 13.35 |
| 74 | 37.00 | 32.50 | 34.00 | 34.63 | 34.66 | 38.49 | 11.89 | 10.31 | 11.51 |
| 75 | 34.00 | 32.00 | 38.00 | 35.45 | 37.40 | 38.70 | 11.30 | 10.24 | 10.18 |
| 77 | 31.00 | 31.67 | 32.33 | 36.70 | 35.84 | 39.97 | 12.06 | 10.65 | 11.19 |
| 78 | 33.00 | 35.33 | 31.00 | 33.41 | 34.31 | 39.41 | 13.00 | 11.00 | 12.83 |
| 79 | - | 43.00 | 41.33 | 36.02 | 36.49 | 42.05 | 11.86 | 9.93 | 8.97 |
| 80 | 32.00 | 28.00 | 26.33 | 35.61 | 37.17 | 39.58 | 13.06 | 11.33 | 11.85 |
| 81 | 29.00 | - | 25.00 | 31.88 | 32.02 | 41.28 | 13.49 | 11.45 | 11.50 |
| 82 | 33.00 | 33.00 | 29.67 | 34.46 | 35.64 | 39.50 | 12.88 | 11.27 | 10.43 |
| 83 | 40.00 | 30.00 | 34.67 | 36.31 | 36.30 | 40.08 | 10.96 | 9.52 | 10.91 |
| 88 | 35.00 | 28.00 | 29.67 | 31.35 | 33.85 | 38.47 | 13.41 | 11.34 | 11.40 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 89 | 35.00 | 40.00 | 41.67 | 32.34 | 36.35 | 40.11 | 11.89 | 9.80 | 10.63 |
| 90 | 40.00 | - | 36.00 | 31.97 | 32.39 | 39.41 | 12.35 | 10.15 | 10.32 |
| 92 | 40.00 | 33.50 | 33.33 | 32.30 | 37.47 | 41.27 | 12.79 | 10.24 | 10.60 |
| 93 | 39.00 | - | 29.00 | 31.41 | 34.21 | 39.19 | 12.15 | 10.82 | 10.90 |
| 94 | 30.00 | 34.33 | 33.67 | 30.14 | 35.66 | 39.02 | 11.45 | 9.64 | 9.86 |
| 95 | 33.00 | 33.33 | 27.33 | 30.16 | 34.73 | 38.72 | 12.65 | 9.61 | 10.60 |
| 97 | 32.00 | 42.00 | 30.00 | 30.64 | 33.54 | 39.27 | 12.39 | 10.95 | 10.25 |
| 99 | 34.00 | 33.67 | 35.33 | 33.21 | 35.29 | 39.68 | 12.44 | 10.23 | 10.23 |
| 101 | 43.00 | - | 32.33 | 31.67 | 34.38 | 33.77 | 13.32 | 9.27 | 9.50 |
| 102 | 25.00 | 29.50 | 29.00 | 30.69 | 32.52 | 38.87 | 11.36 | 9.38 | 9.69 |
| 103 | 30.00 | 33.00 | 31.67 | 33.20 | 35.21 | 37.93 | 11.43 | 10.28 | 11.56 |
| 104 | 38.00 | 34.00 | 34.00 | 33.22 | 36.48 | 36.86 | 12.71 | 12.06 | 14.06 |
| 105 | 36.00 | 33.50 | 28.00 | 32.38 | 32.85 | 36.88 | 10.78 | 10.47 | 10.97 |
| 106 | 36.00 | 35.33 | 31.67 | 29.86 | 32.18 | 34.34 | 12.25 | 10.55 | 11.64 |
| 107 | 36.00 | 40.50 | 33.67 | 31.85 | 36.65 | 39.67 | 11.47 | 9.58 | 9.53 |
| 108 | 36.00 | 37.33 | 33.33 | 30.93 | 35.61 | 39.65 | 11.84 | 9.70 | 9.62 |
| 109 | 37.00 | 31.00 | 30.00 | 31.80 | 32.07 | 37.92 | 11.44 | 10.47 | 10.41 |
| 110 | 34.00 | 43.67 | 34.00 | 33.34 | 35.98 | 41.73 | 10.63 | 9.33 | 9.00 |
| 111 | 33.00 | 32.00 | 34.50 | 32.73 | 34.19 | 39.24 | 12.54 | 11.60 | 10.92 |
| 112 | 29.00 | 31.67 | 30.33 | 30.72 | 32.14 | 35.19 | 14.32 | 11.62 | 11.77 |
| 117 | 30.00 | 29.33 | 28.67 | 32.42 | 34.75 | 37.18 | 13.52 | 11.55 | 11.87 |
| 118 | 35.00 | - | 35.33 | 35.08 | 35.04 | 37.90 | 13.18 | 10.37 | 12.16 |
| 120 | - | 34.67 | 32.33 | 32.33 | 32.11 | 37.53 | 11.41 | 10.56 | 10.94 |
| 121 | 36.00 | 38.00 | 32.33 | 33.43 | 34.18 | 39.55 | 12.40 | 10.94 | 11.16 |
| 122 | 28.00 | 32.00 | 31.67 | 34.81 | 34.36 | 39.11 | 11.29 | 11.21 | 10.86 |
| 123 | 31.00 | 30.33 | 33.00 | 36.57 | 37.61 | 41.06 | 11.92 | 11.02 | 10.79 |
| 124 | 38.00 | 34.00 | 27.33 | 35.90 | 36.35 | 40.20 | 11.99 | 10.57 | 10.49 |
| 125 | 38.00 | 34.33 | 32.67 | 33.43 | 34.16 | 39.39 | 11.69 | 10.75 | 10.84 |
| 126 | 35.00 | 34.67 | 29.67 | 33.16 | 41.59 | 37.06 | 15.08 | 12.51 | 12.48 |
| 127 | 35.00 | 30.67 | 28.67 | 35.19 | 34.57 | 40.32 | 13.11 | 11.28 | 11.84 |
| 128 | 31.00 | 36.00 | 32.67 | 31.98 | 34.05 | 39.18 | 13.80 | 11.06 | 11.60 |
| 130 | 31.00 | 37.50 | 34.33 | 35.58 | 36.32 | 39.18 | 11.78 | 10.75 | 10.82 |
| 131 | 31.00 | 35.33 | 27.67 | 36.06 | 36.40 | 39.53 | 12.92 | 11.50 | 10.05 |
| 132 | 35.00 | 34.00 | 31.50 | 36.56 | 34.98 | 40.41 | 11.01 | 8.62 | 8.49 |
| 133 | 43.00 | 34.50 | 36.00 | 34.80 | 35.59 | 40.55 | 12.19 | 10.64 | 10.35 |
| 134 | 32.00 | 33.00 | 34.33 | 33.28 | 32.78 | 37.66 | 14.70 | 11.69 | 13.22 |
| 135 | 34.00 | 33.00 | 33.33 | 33.02 | 32.77 | 37.06 | 12.97 | 10.95 | 11.78 |
| 136 | 37.00 | 35.00 | 34.00 | 34.49 | 34.30 | 37.54 | 14.05 | 11.70 | 12.37 |
| 137 | 38.00 | 38.67 | 35.33 | 31.92 | 32.60 | 37.95 | 13.79 | 10.69 | 11.74 |
| 138 | 31.00 | 27.00 | 28.67 | 32.56 | 33.60 | 37.77 | 12.69 | 11.27 | 11.33 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 141 | 34.00 | 24.50 | 26.00 | 35.40 | 34.92 | 37.84 | 12.41 | 10.72 | 11.35 |
| 147 | 37.00 | 36.67 | 31.67 | 35.59 | 37.62 | 43.89 | 11.71 | 10.46 | 9.98 |
| 148 | 28.00 | 30.33 | 27.00 | 31.78 | 30.79 | 34.38 | 14.30 | 12.39 | 11.86 |
| 149 | 28.00 | 26.33 | 28.33 | 27.42 | 30.04 | 35.94 | 15.39 | 12.76 | 12.42 |
| 150 | 29.00 | 32.33 | 29.00 | 32.97 | 34.98 | 38.04 | 13.22 | 11.66 | 11.79 |
| 151 | 26.00 | 28.00 | 24.67 | 24.46 | 30.17 | 33.88 | 13.10 | 10.90 | 10.29 |
| 152 | 26.00 | 27.33 | 23.00 | 27.67 | 31.12 | 34.64 | 15.21 | 11.89 | 12.39 |
| 154 | 32.00 | 33.00 | 31.00 | 30.49 | 33.49 | 38.82 | 13.50 | 10.45 | 11.02 |
| 155 | 34.00 | 31.00 | 25.33 | 35.51 | 35.30 | 38.40 | 10.93 | 10.01 | 10.13 |
| 156 | 28.00 | 27.67 | 30.33 | 32.92 | 35.08 | 40.32 | 12.96 | 11.16 | 11.10 |
| 157 | 34.00 | 44.00 | 32.33 | 36.17 | 34.91 | 40.76 | 11.89 | 10.17 | 10.03 |
| 158 | 34.00 | 38.50 | 31.00 | 33.65 | 32.14 | 40.58 | 13.55 | 11.40 | 9.81 |
| 160 | 32.00 | 33.50 | 31.00 | 31.86 | 32.76 | 39.07 | 13.27 | 12.02 | 12.42 |
| 161 | 29.00 | 30.33 | 30.33 | 34.94 | 35.24 | 39.76 | 12.28 | 10.36 | 10.44 |
| 162 | 38.00 | 42.50 | 41.33 | 37.57 | 36.96 | 41.39 | 9.90 | 9.40 | 10.10 |
| 163 | 41.00 | 36.50 | 31.00 | 26.89 | 35.23 | 40.63 | 11.84 | 9.91 | 9.58 |
| 164 | 37.00 | 41.50 | 36.67 | 35.12 | 36.33 | 42.61 | 11.98 | 10.80 | 10.03 |
| 165 | 31.00 | 41.67 | 35.67 | 35.46 | 36.43 | 42.77 | 11.30 | 10.10 | 10.57 |
| 166 | 25.00 | 23.50 | 24.67 | 30.26 | 32.35 | 38.26 | 15.33 | 13.16 | 13.23 |
| 168 | 35.00 | - | 26.50 | 34.01 | 36.08 | 37.48 | 13.90 | 11.78 | 11.05 |
| 169 | 27.00 | 32.67 | 31.33 | 34.53 | 34.75 | 40.05 | 13.10 | 11.21 | 10.77 |
| 170 | 41.00 | 36.00 | 32.00 | 33.42 | 34.92 | 39.61 | 13.19 | 10.53 | 9.80 |
| 175 | 38.00 | 32.50 | 31.67 | 34.10 | 34.94 | 40.19 | 14.64 | 11.28 | 11.98 |
| 176 | 34.00 | - | 32.67 | 32.37 | 34.57 | 40.25 | 13.77 | 12.18 | 11.31 |
| 177 | 28.00 | - | 29.33 | 33.91 | 33.14 | 37.77 | 12.84 | 11.49 | 12.01 |
| 178 | 32.00 | 34.33 | 31.33 | 38.47 | 37.93 | 43.90 | 13.65 | 10.84 | 12.08 |
| 179 | 38.00 | 35.00 | 35.00 | 36.97 | 38.41 | 40.82 | 12.37 | 10.13 | 10.81 |
| 180 | 34.00 | 34.50 | 32.67 | 34.68 | 34.49 | 41.38 | 12.20 | 10.97 | 10.33 |
| 181 | 29.00 | 34.00 | 31.00 | 34.08 | 32.91 | 40.56 | 12.09 | 10.52 | 10.24 |
| 182 | 37.00 | 39.00 | 34.67 | 38.68 | 37.01 | 41.72 | 12.29 | 10.58 | 11.24 |
| 185 | 37.00 | 34.33 | 32.00 | 33.68 | 34.63 | 40.78 | 12.63 | 10.67 | 9.97 |
| 186 | 38.00 | - | 34.67 | 34.86 | 36.20 | 40.68 | 12.25 | 9.89 | 11.40 |
| 187 | 34.00 | 31.67 | 30.33 | 31.95 | 33.78 | 39.74 | 14.03 | 11.08 | 11.75 |
| 188 | 41.00 | 37.00 | 33.00 | 32.61 | 35.63 | 40.83 | 12.33 | 10.44 | 11.21 |
| 189 | 27.00 | 36.00 | 30.00 | 33.52 | 34.82 | 40.95 | 12.60 | 10.83 | 10.42 |
| 190 | 35.00 | 35.67 | 32.00 | 33.35 | 34.48 | 41.01 | 12.21 | 11.10 | 11.36 |
| 191 | 33.00 | 35.50 | 33.00 | 35.50 | 35.85 | 41.76 | 11.15 | 11.58 | 10.46 |
| 192 | 34.00 | 31.00 | 31.00 | 33.96 | 33.99 | 41.09 | 14.28 | 11.63 | 12.54 |
| 193 | 34.00 | 31.00 | 27.33 | 32.22 | 34.23 | 39.51 | 12.33 | 11.07 | 12.76 |
| 194 | 37.00 | 45.67 | 38.33 | 36.25 | 38.78 | 43.75 | 12.26 | 10.70 | 10.56 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 195 | 34.00 | 34.00 | 32.67 | 33.21 | 33.94 | 41.33 | 12.37 | 10.42 | 10.58 |
| 196 | 42.00 | 35.00 | 34.00 | 33.88 | 35.76 | 40.84 | 13.64 | 11.69 | 12.66 |
| 197 | 32.00 | 39.00 | 31.33 | 32.56 | 34.28 | 38.73 | 11.12 | 9.14 | 9.89 |
| 199 | 33.00 | 37.00 | 31.67 | 32.61 | 34.36 | 41.58 | 13.05 | 10.70 | 9.18 |
| 204 | 36.00 | 37.50 | 31.67 | 32.54 | 33.29 | 40.37 | 12.72 | 11.40 | 10.64 |
| 206 | 39.00 | 40.00 | 38.00 | 33.82 | 34.23 | 39.36 | 11.81 | 11.30 | 11.44 |
| 207 | 30.00 | - | 27.33 | 29.97 | 32.35 | 39.06 | 15.05 | 11.75 | 11.43 |
| 208 | 36.00 | - | 35.67 | 34.83 | 35.70 | 41.45 | 11.81 | 10.44 | 11.51 |
| 209 | 30.00 | - | 28.67 | 31.26 | 31.68 | 39.68 | 11.92 | 10.79 | 10.69 |
| 211 | 33.00 | 37.67 | 33.00 | 31.88 | 34.17 | 37.98 | 13.20 | 10.80 | 10.47 |
| 212 | 30.00 | 34.00 | 28.00 | 33.57 | 33.25 | 39.25 | 12.13 | 10.35 | 11.10 |
| 213 | 24.00 | - | 27.00 | 30.77 | 31.54 | 36.63 | 12.22 | 10.30 | 10.23 |
| 214 | 33.00 | 32.00 | 26.33 | 34.89 | 33.05 | 37.34 | 10.27 | 9.56 | 9.00 |
| 216 | 38.00 | - | 39.33 | 33.62 | 35.27 | 40.22 | 12.54 | 10.16 | 10.36 |
| 217 | 39.00 | 44.33 | 36.67 | 33.50 | 35.39 | 40.77 | 12.24 | 10.53 | 9.40 |
| 218 | 30.00 | 34.00 | 33.33 | 31.58 | 34.20 | 39.95 | 12.86 | 11.94 | 10.93 |
| 219 | 35.00 | - | 35.33 | 33.61 | 36.45 | 42.61 | 12.38 | 10.38 | 9.78 |
| 220 | 43.00 | 40.33 | 33.33 | 35.63 | 34.95 | 43.08 | 11.59 | 10.43 | 10.47 |
| 221 | 27.00 | 24.50 | 26.00 | 31.65 | 31.69 | 38.98 | 12.71 | 11.09 | 11.42 |
| 222 | 32.00 | 35.00 | 33.67 | 33.42 | 34.13 | 38.21 | 10.57 | 9.44 | 10.07 |
| 223 | 33.00 | - | 33.00 | 32.94 | 33.42 | 42.39 | 13.56 | 11.34 | 10.63 |
| 224 | 36.00 | - | 31.00 | 33.00 | 36.26 | 41.19 | 14.66 | 10.64 | 10.50 |
| 225 | 31.00 | 28.00 | 26.33 | 36.54 | 36.88 | 42.21 | 12.10 | 10.54 | 10.52 |
| 226 | 29.00 | - | 30.50 | 32.38 | 33.71 | 41.63 | 11.77 | 10.97 | 10.03 |
| MIN | 18.00 | 18.00 | 18.33 | 24.46 | 29.92 | 33.77 | 9.91 | 8.63 | 8.50 |
| MAX | 47.00 | 46.00 | 41.67 | 39.68 | 42.56 | 43.90 | 15.78 | 13.17 | 14.06 |
| AVERAGE | 34.06 | 34.15 | 31.46 | 33.58 | 34.78 | 39.27 | 12.62 | 10.81 | 11.01 |

**Table A-2.** Correlations between lint percent and seed index of the RIL population during 2015 at the St. Joseph, LA, and College Station, and Lubbock, TX environments.

| Correlations | | |
|---|---|---|
| 2015 LA | 2015 CS | 2015 LU |
| -0.4373 | -0.3755 | -0.2735 |